AD-A063 757 CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER --ETC F/G 12/1
GENERAL THEORY OF OPTIMAL ERROR ALGORITHMS AND ANALYTIC COMPLEX--ETC(U)
NOV 78 J F TRAUB, H WOZNIAKOWSKI N00014-76-C-0370
UNCLASSIFIED CMU-CS-78-149 NL

1 OF 2
AD
A063757

2.7

ACCESSION for		
NTIS	White Section	X
DDC	Buff Section	☐
UNANNOUNCED		☐
JUSTIFICATION		
BY		
DISTRIBUTION/AVAILABILITY CODES		
Dist.	AVAIL. and/or SPECIAL	
A		

LEVEL II

12

6

GENERAL THEORY OF OPTIMAL ERROR ALGORITHMS AND ANALYTIC COMPLEXITY.

PART B. ITERATIVE INFORMATION MODEL.

10

J. F. Traub
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

H. Wozniakowski
Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213
(Visiting from the University of Warsaw)

9 Interim rept.,

11 November 1978

12 97p.

14 CMU-CS-78-149

Chapter II was written jointly with B. Kacewicz, University of Warsaw.

This research was supported in part by the National Science Foundation under Grant MCS75-222-55 and the Office of Naval Research under Contract N0014-76-C-0370, NR 033-422.

15 NSF-MCS75-222-55

ABSTRACT

This is the second of a series of papers in which we construct an information based general theory of optimal error algorithms and analytic computational complexity and study applications of the general theory. In our first paper we studied a general information model; here we study an "iterative information" model.

We give a general paradigm, based on the pre-image set of an information operator, for obtaining a lower bound on the error of any algorithm using this information. We show that the order of information provides an upper bound on the order of any algorithm using this information. This upper bound on order leads to a lower bound on the complexity index.

Certain problems are traditionally solved by iteration; others are not. Must this be so? More precisely, for what problems is the class of iterative algorithms empty? For one-point stationary iterations we solve this problem, showing the class is empty unless the problem "index" is finite. We advance a conjecture which characterizes all problems with finite index.

We end the paper with some open problems.

TABLE OF CONTENTS

CHAPTER I

BASIC CONCEPTS

1. INTRODUCTION

This is the second of a series of papers in which we construct an information based general theory of optimal error algorithms and analytic computational complexity and study applications of the general theory. In Traub and Woźniakowski [77c] we studied a general information model; here we study an "iterative information" model. Because of the close connection between this paper and Traub and Woźniakowski [77c] we shall refer to the first paper as Part A.

Although we define all concepts required here, it would probably be helpful to the reader to be familiar with the concepts introduced in Part A. We will sometimes use the same name for a concept here which was defined somewhat differently earlier.

In Part A we showed that given information may not be "strong" enough to solve the problem to within the desired ϵ. It may however turn out that the information can be used to compute a better approximation from a given approximation. By repeating this procedure one may finally be able to solve the problem to within ϵ even for arbitrarily small ϵ. This informal description of iterative information and iterative algorithm will be formalized below.

Certain problems are traditionally solved by iteration; others are not. For example, a zero of a scalar nonlinear function is approximated by iteration using function evaluations; the value of a definite integral of a scalar function is not. Must this be so? More precisely, for what problems is the class of iterative algorithms (in the very general sense of this

paper) empty? For one-point stationary iterations we solve this problem, showing the class is empty unless the problem "index" is finite. We advance a conjecture which characterizes all problems with finite index.

Among the major questions we pose and at least partially answer are:

1. What is a lower bound on the limiting error of any algorithm using given iterative information? See Section 2.
2. In general is there an algorithm which gets close to this lower bound? See Section 2.
3. What is an upper bound on the order of any algorithm using given information? See Section 2.
4. In general is there an algorithm which achieves this bound? See Section 2.
5. What is the most "relevant" information operator for a given problem? See Sections 6 and 8.
6. When is the class of iterative algorithms empty? See Sections 5 and 6.
7. When is the class of iterative algorithms of order m empty? See Section 7.

We summarize major concepts and results of this paper.

<u>Section 2</u>. We define $d(\mathfrak{N},S)$, the <u>limiting diameter of information $\mathfrak{N}$ for the problem S</u> and show (Theorem 2.1) that $\frac{1}{2}d(\mathfrak{N},S)$ provides a best possible lower bound on the limiting error of any algorithm using the information $\mathfrak{N}$. The information $\mathfrak{N}$ is <u>convergent</u> for the problem S iff $d(\mathfrak{N},S) = 0$. We define <u>interpolatory algorithm</u> and show (Theorem 2.2) that any interpolatory algorithm has an error which differs by at most a factor of two from a lower bound on the error. We define $p(\mathfrak{N},S)$, the <u>order of information $\mathfrak{N}$ for the</u>

problem S and prove (Theorem 2.3) that $p(\mathfrak{N},S)$ is an upper bound on the order of any algorithm. This upper bound is achieved (Theorem 2.4) by any interpolatory algorithm.

Section 3. We develop the methodology for complexity analysis and show why it is desirable to minimize the complexity index.

Section 4. In Sections 4-8 we consider linear information operators. The cardinality $\mathrm{card}(\mathfrak{N})$ of a linear information operator is defined and we show (Lemma 4.2) that information operators with finite cardinality equal to n can be represented by n linearly independent linear functionals.

Section 5. We define iterative algorithm and $IT(\mathfrak{N},S)$, the class of all iterative algorithms which use the information operator $\mathfrak{N}$. We prove (Theorem 5.1) that if $d(\mathfrak{N},S)$ is positive then $IT(\mathfrak{N},S)$ is empty.

Section 6. We define index(S), the index of the problem S and prove (Theorem 6.1) that if $\mathrm{card}(\mathfrak{N}) < \mathrm{index}(S)$, then $d(\mathfrak{N},S)$ is positive. We define the basic information operator $\mathfrak{N}^*$ such that $\mathrm{card}(\mathfrak{N}^*) = \mathrm{index}(S)$ and prove (Theorem 6.2) that $\mathfrak{N}^*$ is a convergent information operator. Furthermore if $\mathfrak{N}$ is any convergent information operator then $\mathfrak{N}$ contains $\mathfrak{N}^*$ (Theorem 6.3).

Section 7. We define index(S,m), the mth index of the problem S and prove (Theorem 7.1) that if $\mathrm{card}(\mathfrak{N}) < \mathrm{index}(S,m)$, then $p(\mathfrak{N},S) < m$. We define the basic information operator $\mathfrak{N}_m^*$ such that $\mathrm{card}(\mathfrak{N}_m^*) = \mathrm{index}(S,m)$ and prove (Theorem 7.2) that $p(\mathfrak{N}_m^*,S) \geq m$. We also show that $p(\mathfrak{N},S) \geq m$ implies that $\mathfrak{N}$ contains $\mathfrak{N}_m^*$ (Theorem 7.3).

Section 8. We specify our model of computation for the linear case. We obtain lower and upper bounds on the complexity index and on the mth minimal complexity index.

Section 9. We define information operator with memory and generalize our concepts and theorems on diameter and order of information, interpolatory algorithm, and complexity index.

Section 10. We list some extensions and open problems.

Section 11. We compare the results of the general information model of Part A with the iterative information model of this paper.

Appendix A. We prove (Lemma A.1) that under very weak assumptions, our definition of order of an algorithm agrees with the "classical" definition.

2. DIAMETER AND ORDER OF INFORMATION

As in Part A we consider a linear or nonlinear solution operator S such that

(2.1) $S: \mathfrak{I}_0 \to \mathfrak{I}_2$

where $\mathfrak{I}_0$ is a subset of a linear space $\mathfrak{I}_1$ over the real or complex field and $\mathfrak{I}_2$ is a linear normed space over the real or complex field. We wish to approximate the solution element $\alpha = S(f)$ for all problem elements $f \in \mathfrak{I}_0$. Let x_0 be an initial approximation to the solution α and let ε', $\varepsilon' \in (0,1)$, be a given real number. By solving (or approximating) the problem S we mean that we seek an ε'-approximation $y = y(f)$, $y \in \mathfrak{I}_2$, to α such that

(2.2) $\| y(f)-\alpha\| \le \varepsilon' \| x_0-\alpha\|$.

To find such an approximation we need to know something about the problem element f. Let

(2.3) $\mathfrak{N}: D_{\mathfrak{N}} \subset \mathfrak{I}_1 \times \mathfrak{I}_2 \to \mathfrak{I}_3$

be an iterative information operator (not necessarily linear) where $(f,x) \in D_{\mathfrak{N}}$ for all $f \in \mathfrak{I}_0$ and all x close enough to $\alpha = S(f)$ and where $\mathfrak{I}_3$ is a given space. We call $\mathfrak{N}$ an iterative information operator since we compute $\mathfrak{N}(f,x)$ for different x and the next approximation x_{k+1} is based on the information $\mathfrak{N}(f,x_k)$, $k = 0,1,\ldots$. For brevity $\mathfrak{N}$ will also be called an information operator. The information $\mathfrak{N}$ is stationary and without memory in the sense of Traub [64]. For most problems the information operator $\mathfrak{N}$ is not one-to-one and $\mathfrak{N}(f,x)$ does not uniquely define the solution $\alpha = S(f)$.

This means there exist many different $f \in \mathfrak{J}_0$ with the same information $\mathfrak{N}(f,x)$. Thus for a given f the set of problem elements $\tilde{f}(x)$ such that $\tilde{f}(x) \in \mathfrak{J}_0$ and $\mathfrak{N}(\tilde{f}(x),x) = \mathfrak{N}(f,x)$ determines the "uncertainty" of the information operator $\mathfrak{N}$. Note that $\tilde{f} = \tilde{f}(x)$ is a function of x which has the same information as a problem element f for every x close enough to α. For technical reasons we have to assume that the function $\tilde{f}$ is "regular" at α. To formalize this idea we define equality with respect to $\mathfrak{N}$.

<u>Definition 2.1</u>

We shall say <u>$\tilde{f} = \tilde{f}(x)$ is equal to f with respect to $\mathfrak{N}$</u> iff

(i) $\tilde{f}\colon D_{\tilde{f}} \subset \mathfrak{J}_2 \to \mathfrak{J}_0, \quad \tilde{f} \in W$

where W is a given class and there exists $\Gamma = \Gamma(f) > 0$ such that $J(\Gamma) = \{x\colon \|x-\alpha\| \le \Gamma\} \subset D_{\tilde{f}}$ where $\alpha = S(f)$,

(ii) $\mathfrak{N}(\tilde{f}(x),x) = \mathfrak{N}(f,x), \quad \forall\, x \in D_{\tilde{f}}$.

For brevity we write $\tilde{f} \in V(f)$ where V(f) is the set of all functions $\tilde{f}$ which are equal to f with respect to $\mathfrak{N}$. ∎

The class W describes the regularity of $\tilde{f}$ and its definition depends on the regularity of the solution operator S. We always assume that the constant functions $\tilde{f}(x) \equiv f$ belong to W.

We define the limiting diameter of information as the maximal distance between S operating on two problem elements with the same information at x as x tends to $\alpha = S(f)$. In Section 5 we prove that the limiting diameter of a linear information operator has to be zero in order to solve the problem S iteratively.

Definition 2.2

We shall say $d(\mathfrak{N},S)$ is the limiting diameter of information $\mathfrak{N}$ for the problem S iff

(2.4) $$d(\mathfrak{N},S) = \sup_{f\in\mathfrak{I}_0} \sup_{\tilde{f}_1,\tilde{f}_2\in V(f)} \limsup_{x\to\alpha} \| S(\tilde{f}_1(x))-S(\tilde{f}_2(x)) \| .$$

We shall say the information $\mathfrak{N}$ is convergent for the problem S iff

(2.5) $$d(\mathfrak{N},S) = 0.$$

$\mathfrak{N}$ is called divergent iff $d(\mathfrak{N},S) > 0$. ∎

Note that the limiting diameter of $\mathfrak{N}$ coincides with the diameter of $\mathfrak{N}$ introduced in Part A for information operators independent of x, i.e., $\mathfrak{N}(f,x) \equiv \mathfrak{N}(f)$.

We give a geometrical interpretation of (2.4). Assume that $S(\tilde{f}(x))$ is continuous at α for any $\tilde{f} \in V(f)$. Then

(2.6) $$d(\mathfrak{N},S) = \sup_{f\in\mathfrak{I}_0} \sup_{\tilde{f}_1,\tilde{f}_2\in V(f)} \| S(\tilde{f}_1(\alpha))-S(\tilde{f}_2(\alpha)) \|.$$

Note that $U(f) = \{S(\tilde{f}(\alpha)): \ \tilde{f} \in V(f)\}$ is the set of all solutions $S(\tilde{f}(\alpha))$ which share the same information as f at α. Then (2.6) yields

(2.7) $$d(\mathfrak{N},S) = \sup_{f\in\mathfrak{I}_0} \operatorname{diam}(U(f))$$

where diam(U(f)) denotes the diameter of the set U(f). See Section 2 of Part A. This can be schematized as follows:

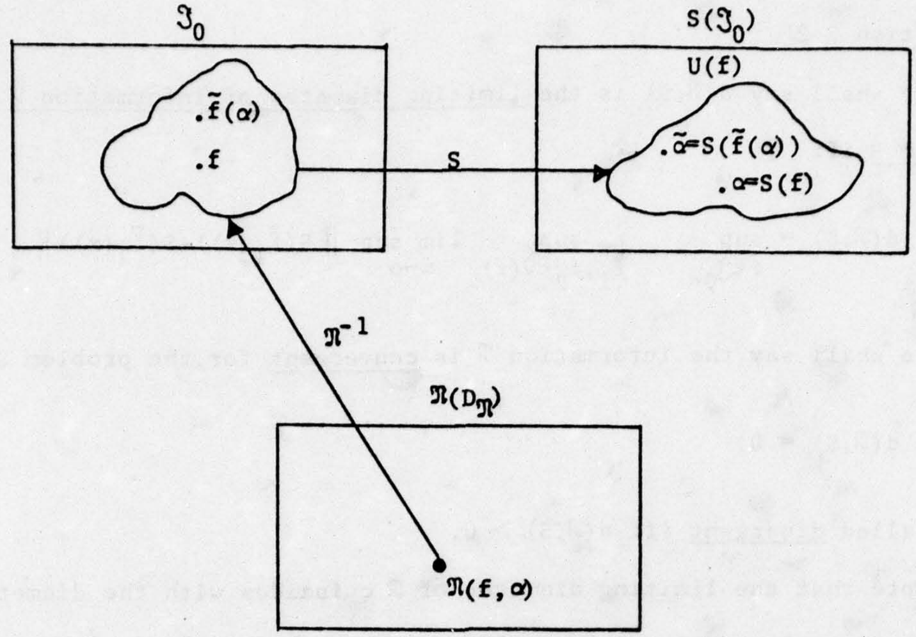

Figure 1

The information $\mathfrak{N}$ is convergent iff the set $U(f)$ contains only one element $\alpha = S(f)$. Thus Figure 1 shows a divergent information operator, i.e., $d(\mathfrak{N},S) > 0$.

We illustrate the concept of limiting diameter of information by an example.

Example 2.1

Let $\mathfrak{J}_1$ be the class of analytic operators f, $f: D_f \subset B_1 \to B_2$ where B_1 and B_2 are Banach spaces and $\dim(B_1) = \dim(B_2)$. Let $\mathfrak{J}_0$ be the class of analytic operators with a unique simple zero, i.e., $f \in \mathfrak{J}_0$ iff $f \in \mathfrak{J}_1$ and there exists a unique $\alpha \in D_f$ such that $f(\alpha) = 0$ and $f'(\alpha)^{-1}$ exists and is bounded. Define

(2.8) $S(f) = f^{-1}(0)$, $\mathfrak{J}_2 = B_2$.

Thus $\alpha = S(f)$ is the solution of the nonlinear equation $f(x) = 0$. Let

(2.9) $\mathfrak{N}(f,x) = [f^{(j)}(x), f^{(j+1)}(x)]$

for a nonnegative integer j, and $f^{(j)}$ denotes the jth Frechet derivative. In this example $\tilde{f}(x) = \tilde{f}(x,\cdot)$ is an analytic operator with respect to the second argument. Let $W = C^{j+1}$ be the class of all functions $\tilde{f}$ which are (j+1) times continuously differentiable at α. An example of $\tilde{f}$ which belongs to V(f) is given by

$$\tilde{f}(x,t) = \begin{cases} f(t) + c & \text{for } j > 0, \\ f(t) + L(t-x)^2 & \text{for } j = 0 \end{cases}$$

where c is a suitably chosen element of $\mathfrak{J}_1$ and L is a bilinear operator. It is easy to verify that

(2.10) $d(\mathfrak{N},S) = \begin{cases} +\infty & \text{for } j > 0, \\ 0 & \text{for } j = 0. \end{cases}$

Thus for $j = 0$ we have convergent information. ■

Remark 2.1

In Part A we defined $r(\mathfrak{N},S)$, the radius of information $\mathfrak{N}$ for the problem S, where $r(\mathfrak{N},S) \in [\frac{1}{2}d(\mathfrak{N},S), d(\mathfrak{N},S)]$. Since our focus in this paper is on convergent information operators, $d(\mathfrak{N},S) = r(\mathfrak{N},S) = 0$, we need not consider the limiting radius of information. ■

We solve the problem (2.2) by an algorithm φ defined as follows. Let

(2.11) $\varphi: D_\varphi \subset \mathfrak{J}_2 \times \mathfrak{N}(D_\mathfrak{N}) \to \mathfrak{J}_2$.

(See also the definition of "permissible algorithm" in Section 3.) Recall that x_0 is an initial approximation to the solution $\alpha = S(f)$. Then the

algorithm φ generates the sequence of approximations by

$$(2.12) \quad x_{i+1} = \varphi(x_i;\ \mathfrak{N}(f,x_i)) \qquad i = 0,1,\ldots\ .$$

Thus, φ is a stationary algorithm and since x_{i+1} depends only on the previously computed approximation x_i, the algorithm φ is without memory in the sense of Traub [64]. In Section 5 we impose some conditions on φ and define the concept of <u>iterative algorithm</u>. Information operators with memory and stationary algorithms with memory are considered in Section 9. We shall not pursue the analysis of nonstationary algorithms in this paper.

Let $\Phi(\mathfrak{N},S)$ be the class of <u>all</u> algorithms defined by (2.11) and (2.12). Let $\varphi \in \Phi(\mathfrak{N},S)$. We examine the convergence of the sequence $\{x_i\}$ to α. Since φ is stationary it suffices to find how x_1 depends on x_0 and whether x_1 converges to α as x_0 tends to α. Recall that the algorithm uses the information $\mathfrak{N}(f,x)$. Suppose that $\tilde{f} \in V(f)$ which means that the information on $\alpha = S(f)$ and $\tilde{\alpha} = S(\tilde{f}(x))$ is exactly the same. Hence any algorithm φ will produce the same approximation to the solution elements α and $\tilde{\alpha}$. Since we are unable to distinguish $f(x)$ from f, an algorithm φ should approximate not only the solution element α but also the solution element $\tilde{\alpha}$. This motivates

<u>Definition 2.3</u>

We shall say $e(\varphi)$ is the <u>limiting error of algorithm φ</u> iff

$$(2.13) \quad e(\varphi) = \sup_{f \in \mathfrak{I}_0} \ \sup_{\tilde{f} \in V(f)} \ \limsup_{x \to \alpha} \|\varphi(x,\mathfrak{N}(f,x)) - S(\tilde{f}(x))\|\ .$$

The algorithm φ is called <u>convergent</u> iff $e(\varphi) = 0$. ■

We are ready to prove that $\frac{1}{2}d(\mathfrak{N},S)$ is a lower bound on $e(\varphi)$ for any algorithm φ from the class $\Phi(\mathfrak{N},S)$.

Theorem 2.1

For any algorithm φ, $\varphi \in \Phi(\mathfrak{N},S)$,

(2.14) $\quad e(\varphi) \geq \frac{1}{2}d(\mathfrak{N},S)$. ∎

Proof

Choose any $\tilde{f}_1$ and $\tilde{f}_2$ from $V(f)$ where $\alpha = S(f)$. Then

$$\overline{\lim_{x\to\alpha}} \| S(\tilde{f}_1(x))-S(\tilde{f}_2(x)) \| \leq \overline{\lim_{x\to\alpha}}(\| \varphi(x,\mathfrak{N}(f,x))-S(\tilde{f}_1(x)) \| + \| \varphi(x,\mathfrak{N}(f,x))- S(\tilde{f}_2(x)) \|) \leq 2\, e(\varphi).$$

Taking the supremum with respect to f and $\tilde{f}_1$, $\tilde{f}_2$ we get from (2.4), $d(\mathfrak{N},S) \leq 2\, e(\varphi)$ which proves (2.14). ∎

Theorem 2.1 states that $\frac{1}{2}d(\mathfrak{N},S)$ is the inherent error of information $\mathfrak{N}$ for any algorithm φ. This is especially interesting for divergent information $\mathfrak{N}$, $d(\mathfrak{N},S) > 0$, since it is then impossible to find an algorithm whose error is less than $\frac{1}{2}d(\mathfrak{N},S)$ no matter how sophisticated an algorithm is used. We showed in Example 2.1 that it is even possible that $d(\mathfrak{N},S) = +\infty$.

We now show that $d(\mathfrak{N},S)$ is an upper bound on "interpolatory algorithms" which are defined as follows.

Definition 2.4

We shall say φ^I, $\varphi^I \in \Phi(\mathfrak{N},S)$, is an <u>interpolatory algorithm</u> iff

(2.15) $\quad \varphi^I(x,\mathfrak{N}(f,x)) = S(\tilde{f}(x))$

for some $\tilde{f} \in V(f)$. ∎

This means that knowing the information $\mathfrak{N}(f,x)$ one finds a problem element $\tilde{f}(x)$ which has the same information as f at x and the next approximation

is the solution of the problem $S(\tilde{f}(x))$. In practice $\tilde{f}(x)$ is chosen to be "simpler" than f. In some cases, an assumption how to choose an unique $\tilde{f}(x)$ is added. Examples of interpolatory algorithms in the sense of this paper include Newton, secant or any $I_{n,s}$ interpolatory algorithms for the solution of nonlinear equations. See Traub [64] and Woźniakowski [74].

Theorem 2.2

For any interpolatory algorithm φ^I, $\varphi^I \in \Phi(\mathfrak{N},S)$,

(2.16) $e(\varphi^I) \le d(\mathfrak{N},S)$.

Proof

Take any $f \in \mathfrak{J}_0$. Then $\varphi^I(x,\mathfrak{N}(f,x)) = S(\tilde{f}_0(x))$ for some $\tilde{f}_0$ such that $\tilde{f}_0 \in V(f)$. Hence

$$\overline{\lim_{x\to\alpha}} \| \varphi^I(x,\mathfrak{N}(f,x)) - S(\tilde{f}(x)) \| = \overline{\lim_{x\to\alpha}} \| S(\tilde{f}_0(x)) - S(\tilde{f}(x)) \| \le d(\mathfrak{N},S)$$

for any $\tilde{f} \in V(f)$. Taking the supremum with respect to f and $\tilde{f}$ we get $e(\varphi^I) \le d(\mathfrak{N},S)$. ■

From Theorems 2.1 and 2.2 we get

Corollary 2.1

There exists a convergent algorithm in $\Phi(\mathfrak{N},S)$ iff the information $\mathfrak{N}$ is convergent for the problem S, i.e., $d(\mathfrak{N},S) = 0$. ■

For convergent information operators, $S(\tilde{f}(x))$ approaches $\alpha = S(f)$ as x tends to α for all $\tilde{f} \in V(f)$. We define the "order of information" which measures the speed of convergence of $S(\tilde{f}(x))$ to $S(f)$ for a worst case. Let A be a set of real numbers defined by

(2.17) $A = \{q\colon q \geq 1, \forall f \in \mathfrak{J}_0, \alpha = S(f)$, and $\forall \tilde{f} \in V(f)$ we have

$$\lim_{x \to \alpha} \frac{\| S(\tilde{f}(x)) - S(f) \|}{\| x - \alpha \|^{q - \eta}} = 0, \quad \forall \eta > 0\}.$$

Definition 2.5

We shall say $p(\mathfrak{N},S)$ is the order of information $\mathfrak{N}$ for the problem S iff

$$(2.18)\quad p(\mathfrak{N},S) = \begin{cases} 0 & \text{if } A \text{ is empty,} \\ \sup A & \text{otherwise.} \end{cases}$$

■

Note that for divergent information operators $p(\mathfrak{N},S) = 0$. It is easy to verify that $p(\mathfrak{N},S)$ is an integer for a sufficiently regular function $S(\tilde{f}(x))$. Roughly speaking, the order $p = p(\mathfrak{N},S)$ measures how fast $S(\tilde{f}(x))$ tends to $S(f)$, $\| S(\tilde{f}(x)) - S(f) \| = O(\| x - \alpha \|^p)$ for all $\tilde{f} \in V(f)$. We prove that any algorithm from $\Phi(\mathfrak{N},S)$ has order no greater than the order of information. This significantly simplifies the complexity analysis since the maximal order of an algorithm is independent of the "structure" of that algorithm, depending only on the information used. See Section 3. We illustrate the concept of the order of information by an example.

Example 2.2

Consider the solution of nonlinear equations defined in Example 2.1. Let

$$(2.19)\quad \mathfrak{N}(f,x) = [f(x), f'(x), \ldots, f^{(n-1)}(x)], \quad n \geq 2,$$

be standard information. Let $W = C^n$ be the class of all functions $\tilde{f}$ which are n times continuously differentiable at α. Then $\tilde{f} \in V(f)$ means $\tilde{f}^{(j)}(x,t)\big|_{t=x} = f^{(j)}(x)$ for $j = 0,1,\ldots,n-1$ where $\tilde{f}^{(j)}$ denotes the jth Frechet derivative with respect to the second argument. Furthermore

$$\tilde{f}(x,t) - f(t) = \int_0^1 \{\tilde{f}^{(n)}(x,\tau t+(1-\tau)x) - f^{(n)}(\tau t+(1-\tau)x)\}(t-x)^n \cdot \frac{(1-\tau)^{n-1}}{(n-1)!}\, d\tau$$

which yields for $t = \alpha$, $\tilde{f}(x,\alpha) = O(\|x-\alpha\|^n)$. Since $0 = \tilde{f}(x,\tilde{\alpha}) = \tilde{f}(x,\alpha) + \tilde{f}'(x,\alpha)(\tilde{\alpha}-\alpha) + O(\|\tilde{\alpha}-\alpha\|^2)$ and $\tilde{f}'(x,\alpha)$ tends to $f'(\alpha)$ which is invertible, we get

$$\|\tilde{\alpha}-\alpha\| = O(\|\tilde{f}(x,\alpha)\|) = O(\|x-\alpha\|^n).$$

This bound is sharp which yields

$$p(\mathfrak{N},S) = n.$$

See also Woźniakowski [75] where the order of information for nonlinear equations was first defined and analyzed. The study of standard information can be found in Traub and Woźniakowski [76c, 77a, 77b]. ■

For convergent information Corollary 2.1 guarantees the existence of convergent algorithms. Let φ be an algorithm from $\Phi(\mathfrak{N},S)$. We want to examine how fast $\varphi(x,\mathfrak{N}(f,x))$ converges to $\alpha = S(f)$ as x tends to α. Let B be a set of real numbers defined by

(2.20) $B = \{q: q \ge 1, \forall f \in \mathfrak{J}_0, \alpha = S(f)$, and $\forall \tilde{f} \in V(f)$ we have

$$\lim_{x\to\alpha} \frac{\|\varphi(x,\mathfrak{N}(f,x)) - S(\tilde{f}(x))\|}{\|x-\alpha\|^{q-\eta}} = 0, \quad \forall \eta > 0\}.$$

<u>Definition 2.6</u>

We shall say $p(\varphi)$ is the <u>order of algorithm φ</u> iff

$$(2.21)\quad p(\varphi) = \begin{cases} 0 & \text{if B is empty} \\ \sup B & \text{otherwise.} \end{cases}$$

■

Note that for a divergent information operator $\mathfrak{N}$, the order of any algorithm φ from $\Phi(\mathfrak{N},S)$ is equal to zero. Definition 2.6 of the order $p(\varphi)$ differs from the "classical" definition of order where $\varphi(x,\mathfrak{N}(f,x))$ is compared only with $\alpha = S(f)$. In Appendix A we show that for all algorithms of practical interest the "classical" order is equal to $p(\varphi)$.

Let $g(x) = \varphi(x,\mathfrak{N}(f,x))-S(\tilde{f}(x))$. Then (2.20) yields that $g^{(j)}(\alpha) = 0$ for $j = 0,1,\ldots,k$ where $k = p(\varphi)-1$ if $p(\varphi)$ is an integer and $k = \lfloor p(\varphi) \rfloor$ otherwise. If $p(\varphi) = +\infty$ and g is analytic in a neighborhood of α then $g(x) \equiv 0$ which means that $\varphi(x,\mathfrak{N}(f,x)) = S(\tilde{f}(x)) = \alpha$ for x close to α. Thus the problem S can be solved in one step. Therefore, we shall say <u>φ is a direct algorithm if $p(\varphi) = +\infty$</u>.

We now prove that the order of information is an upper bound on the order of any algorithm φ and that every interpolatory algorithm achieves this bound.

<u>Theorem 2.3</u>

For any algorithm φ, $\varphi \in \Phi(\mathfrak{N},S)$,

(2.22) $p(\varphi) \le p(\mathfrak{N},S)$. ■

<u>Proof</u>

Suppose first that $B = \emptyset$. Then $p(\varphi) = 0 \le p(\mathfrak{N},S)$. Without loss of generality we can then assume $B \neq \emptyset$. Let q be an arbitrary element of B. Let $f \in \mathfrak{J}_0$ and $\tilde{f}_1,\tilde{f}_2 \in V(f)$. Then

$$\| S(\tilde{f}_1(x))-S(\tilde{f}_2(x)) \| \le \| \varphi(x,\mathfrak{N}(f,x))-S(\tilde{f}_1(x)) \| +$$
$$+ \| \varphi(x,\mathfrak{N}(f,x))-S(\tilde{f}_2(x)) \| = o(\| x-\alpha\|^{q-\eta}), \quad \forall\, \eta > 0.$$

This proves that $q \in A$, see (2.17), and $B \subset A$. Thus $\sup B \le \sup A$ which means $p(\varphi) \le p(\mathfrak{N},S)$. Hence (2.22) is proven. ∎

Theorem 2.3 states that no algorithm can approximate $S(\tilde{f}(x))$ with order higher than the order of information $p(\mathfrak{N},S)$. We show that the bound $p(\mathfrak{N},S)$ is achieved by the order of any interpolatory algorithm.

Theorem 2.4

For any interpolatory algorithm φ^I, $\varphi^I \in \Phi(\mathfrak{N},S)$,

(2.23) $p(\varphi^I) = p(\mathfrak{N},S)$. ∎

Proof

Without loss of generality assume that $A \neq \emptyset$. Let $q \in A$. Then

(2.24) $\| S(\tilde{f}_1(x)) - S(\tilde{f}_2(x)) \| = o(\| x-\alpha \|^{q-\eta}), \quad \forall \eta > 0,$

for any $f \in \mathfrak{J}_0$ and $\tilde{f}_1, \tilde{f}_2 \in V(f)$. Since $\varphi^I(x,\mathfrak{N}(f,x)) = S(\tilde{f}_0(x))$ for $\tilde{f}_0 \in V(f)$, we get from (2.24)

$$\| \varphi^I(x,\mathfrak{N}(f,x)) - S(\tilde{f}(x)) \| = \| S(\tilde{f}_0(x)) - S(\tilde{f}(x)) \| = o(\| x-\alpha \|^{q-\eta})$$

since $\tilde{f}, \tilde{f}_0 \in V(f)$. This proves that $q \in B$ and $A \subset B$. Thus $\sup A \le \sup B$ and $p(\mathfrak{N},S) \le p(\varphi^I)$. From Theorem 2.3 we find $p(\varphi^I) = p(\mathfrak{N},S)$ which completes the proof. ∎

From Theorems 2.3 and 2.4 we obtain

Corollary 2.2

An interpolatory algorithm φ^I achieves the maximal order $p(\mathfrak{N},S)$ in the class $\Phi(\mathfrak{N},S)$,

$$p(\varphi^I) = p(\mathfrak{N},S) = \sup_{\varphi \in \Phi(\mathfrak{N},S)} p(\varphi). \quad ∎$$

The problem of maximal order algorithms was first posed by Traub [61] for nonlinear equations. It was solved for a particular class of non-stationary iterations using standard information for scalar equations by Brent, Winograd, and Wolfe [73]. Theorems 2.3 and 2.4 were established in full generality for nonlinear equations by Woźniakowski [75] and used by many people to establish the maximal order of certain iterative algorithms and/or to compare different information operators from a computational complexity point of view, see e.g., Kacewicz [75, 76a, 76b], Meersman [76a, 76b], Traub and Woźniakowski [76], Wasilkowski [77] and Woźniakowski [72, 74, 76]. Compare also with recent papers of Werschulz [77a, 77b] who uses information operators $\mathfrak{N}(f,h)$ for numerical quadrature and differentiation.

3. COMPLEXITY OF GENERAL INFORMATION

We present our general model of computation; it is quite similar to the model in Part A. We discuss complexity in this model and derive bounds on the "complexity index".

Model of Computation

(i) We assume that the computations are performed on a random access machine. Let p be a primitive operation. Examples of primitive operations include arithmetic operations, the evaluation of a square root or of an integral. Let comp(p) be the complexity (the total cost) of p; comp(p) must be finite. Suppose that P is a given collection of primitives. The choice of P and comp(p), $p \in P$, *are arbitrary and can depend on the particular problem being solved.*

(ii) Let $\mathfrak{N}$ be an information operator. We say that $\mathfrak{N}$ is a permissible information operator with respect to P if $\mathfrak{N}(f,x)$ can be computed by a finite number of primitive operations from P for all (f,x) under consideration. Let comp($\mathfrak{N}(f,x)$) denote the information complexity of computing $\mathfrak{N}(f,x)$. We assume that if $\mathfrak{N}(f,x)$ requires the evaluation of primitives $p_1, p_2, \ldots, p_k$ then

$$\mathrm{comp}(\mathfrak{N}(f,x)) = \sum_{i=1}^{k} \mathrm{comp}(p_i).$$

(iii) Let φ be an algorithm which uses the permissible information $\mathfrak{N}$. To evaluate $\varphi(\mathfrak{N}(f,x))$ we:

(a) compute $y = \mathfrak{N}(f,x)$,

(b) compute $\varphi(x,y)$.

The complexity of computing y is given by (ii). We say that φ is a permissible algorithm with respect to P if $\varphi(x,y)$ can be computed by a finite number of primitive operations from P for all $(x,\mathfrak{N}(f,x))$ under consideration. Let $\text{comp}(\varphi(x,y))$ be the combinatory complexity of computing $\varphi(x,y)$. We assume that if $\varphi(x,y)$ requires the evaluation of primitives $q_1,q_2,\ldots,q_j$ then

$$\text{comp}(\varphi(x,y)) = \sum_{i=1}^{j} \text{comp}(q_i).$$

■

Let $\mathfrak{N}$ be a convergent permissible information operator with order of information $p(\mathfrak{N},S)$ greater than unity. Let φ be a convergent permissible algorithm from the class $\Phi(\mathfrak{N},S)$ with order $p = p(\varphi)$ greater than unity. We analyze the complexity of finding an approximation $y = y(f)$ to the solution $\alpha = S(f)$ for $f \in \mathfrak{J}_0$ using algorithm φ where

$$(3.1)\quad \| y(f)-\alpha\| \le \varepsilon' \| x_0-\alpha\|$$

for a given initial approximation x_0 and a given number $\varepsilon' \in (0,1)$. The analysis is primarily based on Traub and Woźniakowski [76a, 77b] where the nonlinear equation problem is studied.

Assume that φ generates the sequence $x_i = \varphi(x_{i-1},\mathfrak{N}(f,x_{i-1}))$ $i = 1,2,\ldots,k$ such that

$$(3.2)\quad e_i = G_i\, e_{i-1}^{p},\quad e_i = \| x_i-\alpha\|,\quad i = 1,2,\ldots,k$$

where $G_i = G_i(f)$ satisfies

$$(3.3)\quad 0 < \underline{G} \le G_i \le \bar{G} < +\infty$$

and algorithm φ is terminated after k steps. From (3.2) we get

$$(3.4)\quad e_i = \left(\frac{1}{\omega_i}\right)^{p^i-1} e_0 \quad \text{where } \frac{1}{\omega_i} = (G_1^{p^{i-1}} G_2^{p^{i-2}} \cdots \cdot G_i)^{\frac{1}{p^i-1}} e_0.$$

Note that $(e_0\omega_i)^{1-p}$ is the geometric mean of the $G_1, G_2, \ldots, G_i$. Furthermore $e_i < e_0$ iff $\omega_i > 1$. From (3.3) we have

$$(3.5)\quad \frac{1}{\underline{\omega}} = (\underline{G})^{\frac{1}{p-1}} e_0 \le \frac{1}{\omega_i} \le (\bar{G})^{\frac{1}{p-1}} e_0 = \frac{1}{\bar{\omega}}$$

Assume that $\bar{\omega} > 1$. For a given ϵ', let k be the smallest index for which $e_k \le \epsilon' e_0$. Define $\epsilon < \epsilon'$ so that

$$(3.6)\quad e_k = \epsilon e_0.$$

From (3.4) and (3.6) we find

$$(3.7)\quad \left(\frac{1}{\omega_k}\right)^{p^k-1} = \epsilon \text{ and } k = \frac{g(\omega_k)}{\log p}$$

where

$$(3.8)\quad g(\omega) = \log\left(1 + \frac{t}{\log \omega}\right), \quad t = \log 1/\epsilon.$$

We take all logarithms for the remainder of this paper to base 2.

Let comp = comp(φ,f) be the complexity of computing x_k starting at x_0. We do not consider the complexity of finding an initial approximation x_0. See Kung [76] where this problem is considered. The cost of the ith step is equal to comp($\mathfrak{N}(f,x_i)$) + comp($\varphi(x_i, \mathfrak{N}(f,x_i))$). For simplicity we assume that the information complexity and the combinatory complexity do not depend on x_i. Then the cost of each step is equal to $c(\varphi,f)$ = comp($\mathfrak{N}(f,x)$) + comp($\varphi(x,\mathfrak{N}(f,x))$) and comp($\varphi$,f) = kc($\varphi$,f). From (3.7) we get

(3.9) $\quad \text{comp}(\varphi,f) = z\, g(\omega_k)$

where

(3.10) $\quad z = z(\varphi,f) \overset{\text{df}}{=} \dfrac{c(\varphi,f)}{\log\, p(\varphi)}$

is called the complexity index of φ for f. By the complexity index $z(\varphi)$ of an algorithm φ we mean

(3.11) $\quad z(\varphi) = \sup\limits_{f \in \mathfrak{J}_0} z(\varphi,f).$

Remark 3.1

We have been considering the case $p = p(\varphi) > 1$. For completeness we exhibit the case $p = 1$ with the additional assumption that $G_i < 1$, $i = 1,2,\ldots,k$. Then $e_i = \left(\frac{1}{\omega_i}\right)^i e_0$ with $\frac{1}{\omega_i} = (G_1\, G_2 \cdot \ldots \cdot G_i)^{1/i}$. Hence $\frac{1}{\omega_i}$ is the geometric mean of $G_1, G_2, \ldots, G_i$. Assume that $e_k = \varepsilon e_0$. Then $k = (\log 1/\varepsilon)/\log \omega_k$ and the complexity $\text{comp}(\varphi,f)$ of computing x_k is given by

$$\text{comp}(\varphi,f) = z \log \frac{1}{\varepsilon}$$

where $z = z(\varphi,f,\varepsilon) = \dfrac{c(\varphi,f)}{\log \omega_k}$ and $z(\varphi,\varepsilon) = \sup\limits_{f \in \mathfrak{J}_0} z(\varphi,f,\varepsilon)$ is called the complexity index for $p = 1$. Assume that $1/\underline{\omega} \le 1/\omega_i \le 1/\bar{\omega}$ with $\bar{\omega} > 1$. Then

$$\frac{c(\varphi,f)}{\log \underline{\omega}} \le z(\varphi,f,\varepsilon) \le \frac{c(\varphi,f)}{\log \bar{\omega}}\,.$$

We shall not pursue the case $p = 1$ further and shall assume for the remainder of this section that $p > 1$. ■

We analyze the complexity $\text{comp}(\varphi,f)$ defined by (3.9). Since $g(\omega)$ is a monotonically decreasing function, (3.5) yields bounds on the complexity

(3.12) $\quad z\, g(\underline{\omega}) \le \text{comp}(\varphi,f) \le z\, g(\bar{\omega}).$

As $\epsilon \to 0$, $g(\omega) \cong \log t$ and comp $\cong z \log t$. Furthermore if we assume that

(3.13) $2 \le \bar{\omega} \le \underline{\omega} \le t$

then (3.12) becomes

$$z(\log t - \log \log t) \le comp(\varphi,f) \le z \log (1+t).$$

See Theorem 3.1 in Traub and Woźniakowski [76a]. In this case $z \log t$ is a good measure of complexity. However if ω_k in (3.9) is near unity, the g factor dominates and for ϵ fixed, comp $\cong z \log \log \omega_k$. Thus for $\omega_k \cong 1$, the effect of the error coefficients G_i and the initial error e_0 cannot be neglected.

Remark 3.1 *and* (3.12) *show how the complexity* $comp(\varphi,f)$ depends asymptotically on ϵ. Using the Θ-notation of Knuth [76],

$$(3.14) \quad comp(\varphi,f) = \begin{cases} \Theta(\log \frac{1}{\epsilon}) & \text{for } p(\varphi) = 1, \\ \Theta(\log \log \frac{1}{\epsilon}) & \text{for } p(\varphi) > 1. \end{cases}$$

This may be contrasted with Theorem 9.2 of Part A where we prove that the complexity of a linear problem can be an "arbitrary" decreasing function of ϵ.

We want to minimize the complexity of computing x_k, i.e., we want to find a permissible algorithm φ with minimal complexity. Since we do not know the value $g(\omega_k)$ in (3.9) we are not able to minimize complexity. However, if (3.13) holds or ϵ is small enough then the minimal complexity is approximately achieved by an algorithm with minimal complexity index. Therefore we seek an algorithm with the smallest complexity index.

This discussion motivates the following definition.

<u>Definition 3.1</u>

We shall say $z(\mathfrak{N},S)$ is the <u>complexity index of the information $\mathfrak{N}$ for the problem S</u> iff

$$(3.15)\quad z(\mathfrak{N},S) = \inf\{z(\varphi) : \varphi \in \Phi_{perm}(\mathfrak{N},S)\}$$

where $\Phi_{perm}(\mathfrak{N},S)$ is the class of all permissible algorithms.

We shall say φ^{mc}, $\varphi^{mc} \in \Phi_{perm}(\mathfrak{N},S)$, is a <u>minimal complexity index algorithm</u> iff

$$(3.16)\quad z(\varphi^{mc}) = z(\mathfrak{N},S).$$

■

Let

$$(3.17)\quad \mathrm{comp}(\mathfrak{N}) = \sup_{(f,x)\in D_{\mathfrak{N}}} \mathrm{comp}(\mathfrak{N}(f,x))$$

be the <u>information complexity of $\mathfrak{N}$</u>. Every algorithm φ which uses $\mathfrak{N}$ has to perform a certain number of primitive operations to produce the next approximation. More precisely, let

$$(3.18)\quad m(\mathfrak{N},S) = \inf_{\varphi\in\Phi_{perm}(\mathfrak{N},S)} \ \sup_{(f,x)\in D} \mathrm{comp}(\varphi(x,\mathfrak{N}(f,x)))$$

where D, $D \subset D_{\mathfrak{N}}$, denotes "hard problems", i.e., $(f,x) \in D$ iff $\mathrm{comp}(\mathfrak{N}(f,x)) = \mathrm{comp}(\mathfrak{N})$. In general, $m(\mathfrak{N},S)$ depends at least linearly on the total number of "independent pieces" of information $\mathfrak{N}$. See Sections 4 and 8 where the "cardinality" of information $\mathfrak{N}$ is introduced and its influence on the combinatory complexity of φ is discussed.

In Theorems 2.3 and 2.4 we showed that the order $p(\varphi)$ of any algorithm φ from the class $\Phi(\mathfrak{N},S)$ is no larger than the order of information $p(\mathfrak{N},S)$ and

there exist algorithms such that $p(\varphi) = p(\mathfrak{N},S)$. From this and (3.14), (3.17), (3.18) we get a lower bound on the complexity index $z(\mathfrak{N},S)$,

$$(3.19) \quad z(\mathfrak{N},S) \geq \frac{\text{comp}(\mathfrak{N}) + m(\mathfrak{N},S)}{\log p(\mathfrak{N},S)} .$$

Furthermore if there exists a maximal order permissible algorithm φ, $p(\varphi) = p(\mathfrak{N},S)$, such that $\text{comp}(\varphi(x,\mathfrak{N}(f,x)) \ll \text{comp}(\mathfrak{N})$ for all $(f,x) \in D_{\mathfrak{N}}$ then

$$(3.20) \quad z(\mathfrak{N},S) \cong \frac{\text{comp}(\mathfrak{N})}{\log p(\mathfrak{N},S)}.$$

Equations (3.19) and (3.20) motivate our interest in the information complexity $\text{comp}(\mathfrak{N})$ and the order of information $p(\mathfrak{N},S)$.

Suppose that the problem $\alpha = S(f)$ can be solved by the use of different information operators from a given class Ψ. We want to know which information operator is more relevant for the problem $\alpha = S(f)$. This discussion motivates the definition of an information operator with minimal complexity index.

<u>Definition 3.2</u>

We shall say an information operator $\mathfrak{N}_1$ is <u>more relevant</u> than an information operator $\mathfrak{N}_2$ for the problem S iff

$$(3.21) \quad z(\mathfrak{N}_1,S) < z(\mathfrak{N}_2,S).$$

We shall say an information operator $\mathfrak{N}^o$ is <u>optimal in the class</u> Ψ, $\mathfrak{N}^o \in \Psi$, iff

$$(3.22) \quad z(\mathfrak{N}^o,S) = \inf_{\mathfrak{N}\in\Psi} z(\mathfrak{N},S). \qquad \blacksquare$$

In Section 8 we will study optimal information operators for the linear case.

We compare two information operators by their complexity indices and an information operator with minimal complexity index is called optimal. Note however that the complexity of an algorithm using an optimal information operator also depends on error coefficients G_i and the initial approximation. As noted earlier the minimal complexity index can be a poor measure of complexity. See Traub and Woźniakowski [76a] for a discussion of this point.

CHAPTER II

ITERATIVE LINEAR INFORMATION

We assume throughout this chapter that the iterative information operator $\mathfrak{N}$ is a linear operator.

4. CARDINALITY OF LINEAR INFORMATION

Let $\mathfrak{N}$ be an information operator such that

$$(4.1)\quad \mathfrak{N}: \mathfrak{I}_1 \times X \rightarrow \mathfrak{I}_3$$

where X is an open subset of $\mathfrak{I}_2$. We deal with information operators which are linear with respect to the first argument, i.e.,

$$(4.2)\quad \mathfrak{N}(c_1 f_1 + c_2 f_2, x) = c_1 \mathfrak{N}(f_1, x) + c_2 \mathfrak{N}(f_2, x)$$

for any elements f_1, f_2 from $\mathfrak{I}_1$, any constants c_1, c_2 and any $x \in X$. Let

$$(4.3)\quad \ker \mathfrak{N}(\cdot, x) = \{f: \mathfrak{N}(f,x) = 0\}$$

be the kernel of $\mathfrak{N}(\cdot,x)$ for any $x \in X$. As we shall see in Section 5 the kernel of $\mathfrak{N}$ will play an essential role.

Let $\mathfrak{N}_1: \mathfrak{I}_1 \times X \rightarrow \mathfrak{I}_3$ and $\mathfrak{N}_2: \mathfrak{I}_1 \times X \rightarrow \mathfrak{I}_3'$ be two information operators where the space $\mathfrak{I}_3'$ is not necessarily equal to $\mathfrak{I}_3$.

<u>Definition 4.1</u>

We shall say <u>$\mathfrak{N}_1$ is contained in $\mathfrak{N}_2$</u> (briefly $\mathfrak{N}_1 \subset \mathfrak{N}_2$) iff $\ker \mathfrak{N}_2(\cdot,x) \subset \ker \mathfrak{N}_1(\cdot,x)$, $\forall x \in X$.

We shall say <u>$\mathfrak{N}_1$ is equivalent to $\mathfrak{N}_2$</u> (briefly $\mathfrak{N}_1 \asymp \mathfrak{N}_2$) iff $\ker \mathfrak{N}_1(\cdot,x) = \ker \mathfrak{N}_2(\cdot,x)$, $\forall x \in X$. ■

Note that $A = \ker \mathfrak{N}(\cdot,x)$ is a linear subspace of $\mathfrak{J}_1$. Recall that there exists a linear subspace $A^{\perp}$ of $\mathfrak{J}_1$ such that

$$(4.4)\quad \mathfrak{J}_1 = A \oplus A^{\perp}$$

where $A^{\perp}$ is isomorphic to the quotient space $\mathfrak{J}_1/A$ and

$$(4.5)\quad \operatorname{codim} A \overset{\mathrm{df}}{=} \dim A^{\perp} = \dim \mathfrak{J}_1/A.$$

$A^{\perp}$ is called an algebraic complement of A.

To simplify further considerations we assume that the domain X of the information is chosen in such a way that $\operatorname{codim} \ker \mathfrak{N}(\cdot,x) \equiv \text{const}$ for every $x \in X$.

Definition 4.2

We shall say that $\operatorname{card}(\mathfrak{N})$ is the cardinality of the information $\mathfrak{N}$ iff

$$(4.6)\quad \operatorname{card}(\mathfrak{N}) = \operatorname{codim} \ker \mathfrak{N}(\cdot,x),\ x \in X. \qquad \blacksquare$$

As an example consider the information operator $\mathfrak{N}$ defined by

$$(4.7)\quad \mathfrak{N}(f,x) = [L_1(f,x),\ldots,L_n(f,x)]$$

where $L_j\colon \mathfrak{J}_1 \times X \to \mathbb{C}$ is a linear functional with respect to the first argument, $j = 1,2,\ldots,n$. We assume that $L_1(\cdot,x),\ldots,L_n(\cdot,x)$ are linearly independent for every $x \in X$.

Lemma 4.1

Let $\mathfrak{N}$ be defined by (4.7). Then

$$(4.8)\quad \operatorname{card}(\mathfrak{N}) = n \qquad \blacksquare$$

Proof

Note that $A(x) = \ker \mathfrak{N}(\cdot,x) = \{f\colon L_j(f,x) = 0 \text{ for } j = 1,2,\ldots,n\}$. Let $\mathfrak{F}_1 = A(x) \oplus A(x)^{\perp}$ and $\operatorname{lin}(\xi_1,\xi_2,\ldots,\xi_m) \subset A(x)^{\perp}$ for linearly independent $\xi_1,\xi_2,\ldots,\xi_m$. Let $f = \sum_{j=1}^{m} c_j\xi_j$, $f \in A(x)^{\perp}$. We want to find $\vec{c} = [c_1,c_2,\ldots,c_m]^t$ so that $f \in A(x)$. ("t" denotes the transposition of a vector.) Observe that $L_i(f,x) = \sum_{j=1}^{m} c_j\, L_i(\xi_j,x) = 0$ for $i = 1,2,\ldots,n$ is equivalent to the system of homogenous linear equations,

$$\text{(4.9)} \qquad M\vec{c} = \vec{0}$$

where $M = (L_i(\xi_j,x))$

To prove (4.8) assume that $m > n$. Then (4.9) has a nonzero solution $[c_1,c_2,\ldots,c_m]^t$ and $f = \sum_{j=1}^{m} c_j\xi_j \in A(x)^{\perp} \cap A(x)$. *Thus $f = 0$ which contradicts* the linear independence of $\xi_1,\ldots,\xi_m$. Hence $\operatorname{card}(\mathfrak{N}) \le n$. To prove that $\operatorname{card}(\mathfrak{N}) = n$ it suffices to observe that $L_1(\cdot,x),\ldots,L_n(\cdot,x)$ are linearly independent iff $M^t\vec{d} = \vec{0}$ has the unique solution $\vec{d} = \vec{0}$ which holds iff $m = n$. ∎

We now show that any linear information operator $\mathfrak{N}$ may be represented by linear functionals.

Lemma 4.2

Let $\mathfrak{N}$ be a linear information operator and $n = \operatorname{card}(\mathfrak{N}) \le +\infty$. Then there exist $L_1,L_2,\ldots,L_n$ such that

(i) $L_j\colon \mathfrak{F} \times X \to \mathbb{C}$, $j = 1,2,\ldots,n$

(ii) $L_j(\cdot,x),\ldots,L_n(\cdot,x)$ are linearly independent linear functionals

and $\mathfrak{N} \simeq \mathfrak{N}_1$ where $\mathfrak{N}_1 = [L_1,L_2,\ldots,L_n]$ ∎

Proof

Let $\mathfrak{J}_1 = A(x) \oplus A(x)^{\perp}$ where $A(x) = \ker \mathfrak{N}(\cdot,x)$. Then $A(x)^{\perp} = \mathrm{lin}(\xi_1(x),\ldots,\xi_n(x))$ where $\xi_i(x)$ are linearly independent. Every element f, $f \in \mathfrak{J}_1$, has a unique representation

$$f = f_0(x) + \sum_{i=1}^{n} L_i(f,x)\xi_i(x)$$ for some linearly independent functionals L_i such that $L_i(\xi_i(x),x) = \delta_{ij}$ and $f_0(x) \in A(x)$. Then the information operator $\mathfrak{N}_1 = [L_1,L_2,\ldots,L_n]$ satisfies

$$\ker \mathfrak{N}_1(\cdot,x) = \{f: L_i(f,x) = 0,\ i = 1,2,\ldots,n\} = A(x) = \ker \mathfrak{N}(\cdot,x).$$

This proves that $\mathfrak{N} \asymp \mathfrak{N}_1$. ∎

Let $A(\mathfrak{J}_1)$ be the class of all linear subspaces of $\mathfrak{J}_1$. Consider

(4.10) $A: X \subset \mathfrak{J}_2 \to A(\mathfrak{J}_1)$,

i.e., $A(x)$ is a linear subspace of $A(\mathfrak{J}_1)$ for any $x \in X$. We show the relationship between transformations A and linear information operators.

Lemma 4.3

Let $A: X \subset \mathfrak{J}_2 \to A(\mathfrak{J}_1)$ and $\mathrm{codim}\, A(x) = n$, $\forall x \in X$. Then there exists a linear information operator

(4.11) $\mathfrak{N}(f,x) = [L_1(f,x),\ldots,L_n(f,x)]$.

where $L_1(\cdot,x),\ldots,L_n(\cdot,x)$ are linearly independent linear functionals such that

(4.12) $\ker \mathfrak{N}(\cdot,x) = A(x)$. ∎

Proof

Let $\mathfrak{I}_1 = A(x) \oplus A(x)^{\perp}$ where $A(x)^{\perp} = \mathrm{lin}(\xi_1(x),\ldots,\xi_n(x))$ for linearly independent $\xi_1(x),\ldots,\xi_n(x)$. Then every element $f \in \mathfrak{I}_1$ has a unique representation

$$(4.13) \quad f = f_0(x) + \sum_{j=1}^{n} c_j(f,x)\,\xi_j(x)$$

where $f_0(x) \in A(x)$ and $c_j(f,x)$ is a linear functional with respect to f, $c_j(\xi_i(x),x) = \delta_{ij}$. Define

$$(4.14) \quad L_j(f,x) = c_j(f,x) \quad \text{for } j = 1,2,\ldots,n.$$

It is obvious that $L_1(\cdot,x),\ldots,L_n(\cdot,x)$ are linearly independent and

$$\ker \mathfrak{N}(\cdot,x) = \{f\colon L_j(f,x) = 0,\ j = 1,2,\ldots,n\} = A(x),$$

which completes the proof. ■

In the following sections we will consider "regular" linear information operators defined as follows.

Let $\tilde{f}\colon D_{\tilde{f}} \subset \mathfrak{I}_2 \to \mathfrak{I}_0$ where $\mathfrak{I}_0 \subset \mathfrak{I}_1$. Assume that $\mathfrak{I}_1$ is a linear normed space. We shall say that $\tilde{f}$ belongs to the class Lip(k), $k \geq 0$, iff the kth Frechet derivative of $\tilde{f}$ exists at every solution element $\alpha \in S(\mathfrak{I}_0)$ and satisfies a Lipschitz condition. That is, $\tilde{f} \in \mathrm{Lip}(k)$ iff for any $\alpha \in S(\mathfrak{I}_0)$ there exist $\Gamma = \Gamma(\alpha,\tilde{f}) > 0$ and $q = q(\alpha,\tilde{f})$ such that

$$(4.15) \quad \| \tilde{f}^{(k)}(x_1) - \tilde{f}^{(k)}(x_2) \| \leq q \| x_1 - x_2 \|$$

for $\| x_1 - x_2 \| \leq \Gamma$.

Let $L\colon \mathfrak{I}_1 \times X \to \mathbb{C}$ be a linear functional with respect to the first argument. We shall say that $L \in \mathrm{Lip}(k)$ iff $L(h,\cdot) \in \mathrm{Lip}(k)$ for any $h \in \mathfrak{I}_1$. We are ready to define what we mean by $\mathfrak{N} \in \mathrm{Lip}(k)$.

Definition 4.3

We shall say a linear information operator $\mathfrak{N}$ belongs to the class Lip(k) (briefly $\mathfrak{N} \in \mathrm{Lip}(k)$) iff there exist linearly independent linear functionals $L_j: \mathfrak{J}_1 \times X \to \mathbb{C}$, $j = 1,2,\ldots,n = \mathrm{card}(\mathfrak{N})$ such that

(i) $\ker \mathfrak{N}(\cdot,x) = \{h: L_j(h,x) = 0,\ j = 1,2,\ldots,n\}, \quad \forall x \in X,$

(ii) $L_j \in \mathrm{Lip}(k), \quad j = 1,2,\ldots,n.$ ∎

Thus, $\mathfrak{N} \in \mathrm{Lip}(k)$ means that the linear functionals which form the kernel of $\mathfrak{N}$ are k times differentiable and the kth derivative satisfies a Lipschitz condition.

Lemma 4.4

Assume that $\mathfrak{N}: \mathfrak{J}_1 \times X \to \mathfrak{J}_3$ belongs to Lip(k) and $n = \mathrm{card}(\mathfrak{N})$. Let $g_0 \in \ker \mathfrak{N}(\cdot,\alpha)$. Then there exists a function $h: X \to \mathfrak{J}_1$, such that

(i) $\mathfrak{N}(h(x),x) = 0, \quad \forall x \in X,$

(ii) $h \in \mathrm{Lip}(k),$

(iii) $h(\alpha) = g_0.$ ∎

Proof

From Definition 4.3 we have

$$(4.16) \quad \ker \mathfrak{N}(\cdot,x) = \{g: L_j(g,x) = 0,\ j = 1,2,\ldots,n\},\ x \in X,$$

where $L_j \in \mathrm{Lip}(k)$ and $L_1(\cdot,x),\ldots,L_n(\cdot,x)$ are linearly independent linear functionals. Let $\xi_1(x),\ldots,\xi_n(x)$ be elements of $\mathfrak{J}_1$ such that $L_j(\xi_i(x),x) = \delta_{ij}$ for $i,j = 1,2,\ldots,n$ and $\xi_j \in \mathrm{Lip}(k)$. Define

$$(4.17)\quad h(x) = g_0 - \sum_{i=1}^{n} L_i(g_0,x)\,\xi_i(x).$$

Note that $L_j(h(x),x) = L_j(g_0,x) - \sum_{i=1}^{n} L_i(g_0,x)\,\delta_{ij} = 0$ which means that $h(x) \in \ker \mathfrak{N}(\cdot,x)$, $\forall\, x \in X$. Clearly $h \in \mathrm{Lip}(k)$ and $h(\alpha) = g_0$ since $g_0 \in \ker \mathfrak{N}(\cdot,\alpha)$ means $L_i(g_0,\alpha) = 0$ for $i = 1,2,\ldots,n$. This completes the proof. ■

5. WHEN IS THE CLASS OF ITERATIVE ALGORITHMS EMPTY?

Recall that the solution operator S transforms $\mathfrak{I}_0$, $\mathfrak{I}_0 \subset \mathfrak{I}_1$, into $\mathfrak{I}_2$. Throughout the rest of this paper we assume that

(5.1) $\mathfrak{I}_1$ is a linear <u>normed</u> space,

(5.2) $\mathfrak{I}_0$ is <u>open</u>, i.e., for any $f \in \mathfrak{I}_0$ there exists a positive number $\delta = \delta(f)$ such that $f+h \in \mathfrak{I}_0$ for any $h \in \mathfrak{I}_1$ with $\|h\| \le \delta$,

(5.3) $\mathfrak{I}_2$ is a <u>Banach</u> space,

(5.4) S is a Lipschitz operator at every $f \in \mathfrak{I}_0$, i.e., $\|S(f_1)-S(f_2)\| \le q(S,f)\|f_1-f_2\|$ for all f_1 and f_2 sufficiently close to f.

Let $\mathfrak{N}$ be an information operator. Recall that $\tilde{f} : D_f \to \mathfrak{I}_0$ belongs to V(f) and $V(f) \subset W$. See Definition 2.1. We assume that W is the class of functions $\tilde{f}$ such that

(5.5) $\|\tilde{f}(x)-f\| < \delta(f)$,

(5.6) $\tilde{f}$ is a Lipschitz function at $\alpha = S(f)$, i.e., $\|\tilde{f}(x_1)-\tilde{f}(x_2)\| \le q(\tilde{f})\|x_1-x_2\|$

where x, x_1 and x_2 belong to the ball $J = \{x: \|x-\alpha\| \le \Gamma\}$ for a sufficiently small positive $\Gamma = \Gamma(\tilde{f})$.

<u>Remark 5.1</u>

Since $\mathfrak{N}$ is linear, $\mathfrak{N}(\tilde{f}(x)-f,x) = 0$. This means that $h(x) = \tilde{f}(x)-f$ belongs to ker $\mathfrak{N}(\cdot,x)$. Furthermore $\|\tilde{f}(x_1)-\tilde{f}(x_2)\| = \|h(x_1)-h(x_2)\|$ and

(5.5), (5.6) state that $\|h(x)\| < \delta$ and h is a Lipschitz function at α. From Lemma 4.2 we know that $h(x) \in \ker \mathfrak{N}(\cdot,x)$ is equivalent to $L_i(h(x),x) = 0$ for $i = 1,2,\ldots,n = \mathrm{card}(\mathfrak{N})$ for some linear functionals $L_1(\cdot,x),\ldots,L_n(\cdot,x)$. Thus $h = h(x)$ has *to satisfy* n homogenous equations. ∎

We deal with "iterative" algorithms which are defined as follows. Recall that φ is an algorithm, see (2.3), if $\varphi\colon D_\varphi \subset \mathfrak{J}_2 \times \mathfrak{N}(D_{\mathfrak{N}}) \to \mathfrak{J}_2$.

Definition 5.1

We shall say φ is an iterative algorithm (or briefly an iteration) iff for any $f \in \mathfrak{J}_0$, $\alpha = S(f)$, there exists a positive number $\Gamma = \Gamma(f)$ such that for every $x_0 \in J = \{x\colon \|x-\alpha\| \le \Gamma\}$ the sequence $x_{i+1} = \varphi(x_i,\mathfrak{N}(f,x_i))$ is well-defined and

(5.7) $\lim_{i\to\infty} x_i = \alpha$,

(5.8) $\alpha = \varphi(\alpha,\mathfrak{N}(f,\alpha))$. ∎

The conditions (5.7) and (5.8) mean that the algorithm φ produces convergent sequences whenever an initial approximation belongs to the ball J and the solution α is a fixed point of φ. Let $IT(\mathfrak{N},S)$ be the class of all iterative algorithms which use the information operator $\mathfrak{N}$. For which information operators $\mathfrak{N}$ is the class $IT(\mathfrak{N},S)$ non-empty? We show that this problem is related to the limiting diameter of information $d(\mathfrak{N},S)$ defined by (2.4). We need the following Lemma.

Lemma 5.1

Assume there exists $f \in \mathfrak{J}_0$, $\alpha = S(f)$, such that for any $\Gamma > 0$ one can find $f_0 \in \mathfrak{J}_0$ satisfying

(i) $0 < \|\alpha_0 - \alpha\| \le \Gamma$ where $\alpha_0 = S(f_0)$,

(ii) $\mathfrak{N}(f_0, \alpha_0) = \mathfrak{N}(f, \alpha_0)$.

Then $IT(\mathfrak{N}, S) = \emptyset$. ■

Proof

Suppose to the contrary that $IT(\mathfrak{N}, S) \neq \emptyset$ and let $\varphi \in IT(\mathfrak{N}, S)$. Then for f, $\alpha = S(f)$, there exists $\Gamma = \Gamma(f) > 0$ such that $x_{i+1} = \varphi(x_i, \mathfrak{N}(f, x_i))$ is convergent to α for all $\|x_0 - \alpha\| \le \Gamma$. Let f_0 satisfy the assumptions of Lemma 5.1. Apply the iterative algorithm φ to f_0 with the initial approximation $x_0 = \alpha_0$. Since $\alpha_0 = S(f_0)$ is a fixed point of φ, we get $\varphi(\alpha_0, \mathfrak{N}(f_0, \alpha_0)) = \alpha_0$. But $\|\alpha_0 - \alpha\| \le \Gamma$ which means that α_0 can be used as an initial approximation of α. However $x_1 = \varphi(\alpha_0, \mathfrak{N}(f, \alpha_0)) = \varphi(\alpha_0, \mathfrak{N}(f_0, \alpha_0)) = \alpha_0$ which yields $x_{i+1} \equiv \alpha_0 \neq \alpha$. This contradicts that $\{x_i\}$ tends to α. ■

Lemma 5.1 states that if one can find a problem element f_0 which shares the same information as f and α_0 is sufficiently close to but different from α, then the class of iterations $IT(\mathfrak{N}, S)$ is empty. Compare with Theorem 4.1 in Kung and Traub [76b] and Lemma 3.2 in Woźniakowski [76] where a similar proof technique is used.

We are ready to prove

Theorem 5.1

Suppose that (5.1) to (5.6) hold. Let $\mathfrak{N}$ be a linear information operator. If $d(\mathfrak{N}, S) > 0$ then $IT(\mathfrak{N}, S) = \emptyset$. ■

Proof

Since $d(\mathfrak{N},S) > 0$ there exist $f \in \mathfrak{F}_0$, $\alpha = S(f)$, and $\tilde{f} \in V(f)$ such that

(5.9) $S(\tilde{f}(\alpha)) \neq S(f) = \alpha$

Let $h(x) = \tilde{f}(x) - f$ for x close to α. Then $h(x) \in \ker \mathfrak{N}(\cdot,x)$ and due to (5.6) $\tilde{f}$ is a Lipschitz function at α,

$$\| h(x_1)-h(x_2) \| = \| \tilde{f}(x_1)-\tilde{f}(x_2) \| \le q(\tilde{f}) \| x_1-x_2 \|.$$

Due to (5.2) and (5.5), $f + ch(x) \in \mathfrak{F}_0$ for $c \in [0,1]$ since $\| ch(x) \| < \delta(f)$. Consider $g(c) \overset{df}{=} S(f+ch(\alpha))-S(f)$ for $c \in [0,1]$. From (5.4) it follows that g is continuous. Note that $g(0) = 0$ and $g(1) \neq 0$ due to (5.9). Choose $c_0 \in [0,1)$ such that $g(c_0) = 0$ and $g(c_0+\delta) \neq 0$, for $\delta \in (0,1-c_0]$. Let $f_1 = f + c_0 h(\alpha)$. Note that $f_1 \in \mathfrak{F}_0$ and $f_1 + ch(x) \in \mathfrak{F}_0$ for $c \in [0,\delta(f_1)/\delta(f)]$. For such c define $F(x) = S(f_1+ch(x))$. Then $S(f_1) = \alpha$ and $F(\alpha) = S(f_1+ch(\alpha)) \neq \alpha$ for small positive c. We consider the equation $x = F(x)$ for $x \in J = \{x: \|x-\alpha\| \le \Gamma\}$ for small positive Γ. Note that $\|F(x)-\alpha\| = \|S(f_1+ch(x))-S(f_1)\| \le cq(S,f_1) \sup_{x\in J} \| h(x) \| \le \Gamma$ for small c and Γ. Furthermore $\| F(x_1)-F(x_2) \| \le cq(S,f_1) \| h(x_1)-h(x_2) \| \le cq(S,f_1)\, q(\tilde{f}) \| x_1-x_2 \|$. Thus for small c, F is a contraction mapping in J. Since J is a closed ball in the Banach space $\mathfrak{F}_2$, there exist $\alpha_0 \in J$ such that $\alpha_0 = F(\alpha_0)$, i.e., $\alpha_0 = S(f_1+ch(\alpha_0))$ and $\alpha_0 \neq \alpha$. Let $f_0 = f_1 + ch(\alpha_0)$. Since $h(\alpha_0) \in \ker \mathfrak{N}(\cdot,\alpha_0)$, $\mathfrak{N}(f_0,\alpha_0) = \mathfrak{N}(f_1,\alpha_0)$. Applying Lemma 5.1 for f_1 and f_0 we get $IT(\mathfrak{N},S) = \emptyset$. ∎

Theorem 5.1 states that $d(\mathfrak{N},S) = 0$ is a necessary condition for the class of iterations $IT(\mathfrak{N},S)$ not to be empty. We prove in the next section that unless the cardinality of the information operator $\mathfrak{N}$ is sufficiently large, $d(\mathfrak{N},S) > 0$.

Assume then that $\mathfrak{N}$ is a convergent linear information operator and let $p(\mathfrak{N},S)$ be the order of information $\mathfrak{N}$. Since S is a Lipshitz operator and W is the class of Lipschitz functions, $p(\mathfrak{N},S) \ge 1$.

Lemma 5.2

If $p(\mathfrak{N},S) > 1$ then $IT(\mathfrak{N},S)$ is non-empty. ■

Proof

Let φ^I, $\varphi^I \in \Phi(\mathfrak{N},S)$, be an interpolatory algorithm. From Theorem 2.4 we get $p(\varphi^I) = p(\mathfrak{N},S)$. This yields

(5.10) $\| \varphi^I(x,\mathfrak{N}(f,x))-\alpha\| = o(\| x-\alpha\|^{p-\eta}),\ \forall\eta > 0,$

for all x sufficiently close to $\alpha = S(f)$ with $p = p(\mathfrak{N},S)$. Set $\eta = (p-1)/2$ and $q = (p+1)/2$. Then there exist positive constants C and Γ such that (5.10) can be rewritten

$$\| \varphi^I(x,\mathfrak{N}(f,x))-\alpha\| \le C\| x-\alpha\|^q, \text{ for } \| x-\alpha\| \le \Gamma.$$

Let $e_i = \| x_i-\alpha\|$ where $x_{i+1} = \varphi^I(x_i,\mathfrak{N}(f,x_i))$. Then $e_{i+1} \le Ce_i^q$. Since $q > 1$, then for $e_0 < C^{1/(1-q)}$ the sequence $\{x_i\}$ is convergent to α. Of course, $\alpha = \varphi(\alpha,\mathfrak{N}(f,\alpha))$. This means that φ^I is an iterative algorithm in the sense of Definition 5.1 and $IT(\mathfrak{N},S)$ is non-empty. ■

From the proof of Lemma 5.2 it easily follows

Corollary 5.1

If $\varphi \in \Phi(\mathfrak{N},S)$ and $p(\varphi) > 1$ then φ is an iterative algorithm. ■

For $p(\mathfrak{N},S) = 1$, the class of iterations $IT(\mathfrak{N},S)$ is non-empty if S is a contraction mapping, more precisely if $\| S(\tilde{f}(x_1))-S(\tilde{f}(x_2))\| \le q(f)\| x_1-x_2\|$ for all x_1 and x_2 close enough to $\alpha = S(f)$ and for all $\tilde{f}_1,\tilde{f}_2 \in V(f)$ where $q(f) < 1$. For $p(\mathfrak{N},S) = 1$ and for non-contraction mappings S it seems plausible to conjecture that $IT(\mathfrak{N},S) = \emptyset$. We do not pursue this problem here.

6. INDEX OF THE PROBLEM S

Recall that S: $\mathfrak{I}_0 \to \mathfrak{I}_2$ where $\mathfrak{I}_0$ is a subset of a linear space $\mathfrak{I}_1$. Let $\alpha \in S(\mathfrak{I}_0)$. Define

$$(6.1)\quad G(\alpha) = \{h \in \mathfrak{I}_1 : \forall f \in \mathfrak{I}_0,\ S(f) = \alpha, \text{ we have } S(f+ch) = S(f),\ \ \forall\ |c| < \delta(f)/\|h\|\}.$$

That is, $G(\alpha)$ is a set of elements h which multiplied by a small c and added to f do not change the solution element $\alpha = S(f)$. Due to assumption (5.2), $f+ch \in \mathfrak{I}_0$ for $|c| < \delta(f)/\|h\|$ and therefore $S(f+ch)$ is well-defined. Note that $G(\alpha)$ is a homogenous set, i.e., $h \in G(\alpha)$ implies $ah \in G(\alpha)$ for any constant a. If S is analytic at any $f \in \mathfrak{I}_0$ then

$$(6.2)\quad G(\alpha) = \{h \in \mathfrak{I}_1 : \forall\ f \in \mathfrak{I}_0,\ S(f) = \alpha,\ S^{(j)}(f)h^j = 0, \text{ for } j \geq 1\}.$$

Let $A(\mathfrak{I}_1)$ be the class of all linear subspaces of $\mathfrak{I}_1$.

<u>Definition 6.1</u>

We shall say that index(S) is the <u>index of the problem S</u> iff

$$(6.3)\quad \text{index}(S) = \max_{\alpha \in S(\mathfrak{I}_0)}\ \min_{A \subset G(\alpha) \text{ and } A \in A(\mathfrak{I}_1)} \text{codim } A. \qquad \blacksquare$$

Let $A(\alpha)$ be a subset of $G(\alpha)$ such that $A(\alpha)$ is a linear subspace of $\mathfrak{I}_1$ and has minimal codimension, i.e., $A(\alpha) \subset G(\alpha)$, $A(\alpha) \in A(\mathfrak{I}_1)$ and

$$(6.4)\quad \text{codim } A(\alpha) = \min_{A \subset G(\alpha) \text{ and } A \in A(\mathfrak{I}_1)} \text{codim } A.$$

If $G(\alpha)$ is linear then clearly $A(\alpha) = G(\alpha)$. The index of the problem S is the maximal codimension of $A(\alpha)$. If $G(\alpha)$ is linear for all $a \in S(\mathfrak{I}_0)$ then (6.3) becomes

(6.5) $\text{index}(S) = \max\limits_{\alpha \in S(\mathfrak{J}_0)} \text{codim } G(\alpha).$

In Section 5 we showed that if the limiting diameter of information $d(\mathfrak{N},S)$ is non-zero, then the class of iterations $IT(\mathfrak{N},S)$ is empty. We now prove that $d(\mathfrak{N},S) = 0$ implies that the cardinality of $\mathfrak{N}$ is at least equal to the index S. Recall that throughout this paper we assume that (5.1) to (5.6) hold. We also assume that $S(\mathfrak{J}_0)$ is an open subset of $\mathfrak{J}_2$.

__Theorem 6.1__

Let $\mathfrak{N}: \mathfrak{J}_1 \times S(\mathfrak{J}_0) \to \mathfrak{J}_3$ be an arbitrary linear information operator such that $\mathfrak{N} \in \text{Lip}(0)$ and $\text{card}(\mathfrak{N}) < \text{index}(S)$. Then

(6.6) $d(\mathfrak{N},S) > 0.$ ■

__Proof__

Assume that there exists an information operator $\mathfrak{N}$ such that $\mathfrak{N} \in \text{Lip}(0)$, $n = \text{card}(\mathfrak{N}) < \text{index}(S)$ and $d(\mathfrak{N},S) = 0$. This means that for any f and $\tilde{f}$ such that $\tilde{f} \in V(f)$, $\| S(\tilde{f}(x)) - \alpha \| \to 0$ as x tends to $\alpha = S(f)$. Since $\tilde{f}$ and S are continuous, $S(\tilde{f}(\alpha)) = \alpha$. The information $\mathfrak{N}$ is linear and belongs to Lip(0). This yields that $h(x) = \tilde{f}(x) - f$ is an element of $\ker \mathfrak{N}(\cdot,x) = \{h: L_j(h,x) = 0, j = 1,2,\ldots,n\}$ where the linear functionals $L_j \in \text{Lip}(0)$. Take $\alpha \in S(\mathfrak{J}_0)$ such that $\text{codim ker } A(\alpha) = \text{index}(S)$. Let g_0 be an arbitrary element of $\ker \mathfrak{N}(\cdot,\alpha)$. From Lemma 4.4 it follows that there exists a function $h = h(x)$ such that $\mathfrak{N}(h(x),x) = 0$, $h \in \text{Lip}(0)$ and $h(\alpha) = g_0$. Define $\tilde{f}(x) = f + ch(x)$ for $|c| < \delta(f)/\|g_0\|$ where $S(f) = \alpha$. Then $\tilde{f} \in V(f)$ and $\tilde{f}(\alpha) = f + cg_0$. Thus $S(f + cg_0) = S(f)$ which yields that $g_0 \in G(\alpha)$ and $\ker \mathfrak{N}(\cdot,\alpha) \subset G(\alpha)$. From (6.4) we get $n = \text{codim ker } \mathfrak{N}(\cdot,\alpha) \geq \text{codim } A(\alpha) = \text{index}(S)$. This is a contradiction. Hence $d(\mathfrak{N},S) > 0$ which completes the proof. ■

Theorem 6.1 states that the limiting diameter of information $d(\mathfrak{N},S)$ is nonzero for any linear information operator with cardinality less than index(S). From Theorems 5.1 and 6.1 we get

Corollary 6.1

The class of iterative algorithms $IT(\mathfrak{N},S)$ is empty for any linear information operator $\mathfrak{N}$ such that $\mathfrak{N} \in Lip(0)$ and $card(\mathfrak{N}) < index(S)$. ∎

We now show that there exists a linear information operator $\mathfrak{N}$ such that $\mathfrak{N} \in Lip(0)$, $card(\mathfrak{N}) = index(S)$ and $d(\mathfrak{N},S) = 0$. Recall that $A(\alpha)$ is defined by (6.4). We assume that the domain $\mathfrak{I}_0$ of S is chosen in such a way that codim $A(\alpha)$ does not change for $\alpha \in S(\mathfrak{I}_0)$, i.e., index(S) = codim $A(\alpha)$, $\forall\ \alpha \in S(\mathfrak{I}_0)$. We apply Lemma 4.3 for $A(x)$ where $x \in X = S(\mathfrak{I}_0)$. This yields a linear information operator

$$(6.7)\quad \mathfrak{N}^*(f,x) = [L_1^*(f,x),\ldots,L_{n^*}^*(f,x)]$$

such that $L_1^*(\cdot,x),\ldots,L_{n^*}^*(\cdot,x)$ are linearly independent linear functionals for every x, $n^* = card(\mathfrak{N}) = index(S)$ and $\ker \mathfrak{N}^*(\cdot,x) = A(x)$, $\forall\ x \in S(\mathfrak{I}_0)$. We shall call $\mathfrak{N}^*$ a <u>basic linear information operator</u>.

Definition 6.2

We shall say a <u>solution operator S belongs to the class Lip(0)</u> (briefly $S \in Lip(0)$) iff $\mathfrak{N}^* \in Lip(0)$ where $\mathfrak{N}^*$ is defined by (6.7). ∎

Compare with Definition 4.3. Thus, $S \in Lip(0)$ means that linear functionals whose kernels form $A(x)$ satisfy a Lipschitz condition.

Theorem 6.2

Assume that $S \in \mathrm{Lip}(0)$. Then $\mathfrak{N}^*$ defined by (6.7) is a convergent linear information operator, i.e., $d(\mathfrak{N}^*, S) = 0$. ■

Proof

Take any $f \in \mathfrak{F}_0$, $\alpha = S(f)$. Let $\tilde{f}$ be any function, $\tilde{f} \in \mathrm{Lip}(0)$. such that $\tilde{f} \in V(f)$. Let $h(x) = \tilde{f}(x) - f$. Then $\|h(x)\| < \delta(f)$ and $h(x)$ belongs to $\ker \mathfrak{N}^*(\cdot, x) = A(x)$. Thus $\overline{\lim_{x\to\alpha}}\|S(\tilde{f}(x)) - S(f)\| = \|S(f+h(\alpha)) - S(f)\|$. Since $h(\alpha) \in A(\alpha)$ and $A(\alpha) \subset G(\alpha)$ we get $S(f+ch(\alpha)) = S(f)$ for $|c| < \delta(f)/\|h(\alpha)\|$. But $\|h(\alpha)\| < \delta(f)$ and setting $c = 1$ we get $S(\tilde{f}(\alpha)) = S(f)$. This proves that $d(\mathfrak{N}^*, S) = 0$ which means that $\mathfrak{N}^*$ is convergent. ■

Theorem 6.1 and 6.2 state that it is necessary and sufficient to use linear information operators with cardinality at least equal to the index of the problem S if we wish to compute α by an iterative algorithm.

We proved that the basic linear information operator $\mathfrak{N}^*$ is convergent. We now show that provided a certain technical assumption holds any convergent linear information operator contains $\mathfrak{N}^*$. This means that the information $\mathfrak{N}^*$ must be computed. See Definition 4.1.

Theorem 6.3

Let $G(\alpha)$ be a linear set for all $\alpha \in S(\mathfrak{F}_0)$ and let $\mathfrak{N}$ be a linear information operator such that $\mathfrak{N} \in \mathrm{Lip}(0)$. Then $d(\mathfrak{N}, S) = 0$ implies $\mathfrak{N}^* \subset \mathfrak{N}$.

Proof

Since $G(\alpha)$ is linear, $\ker \mathfrak{N}^*(\cdot, \alpha) = G(\alpha)$, $\forall\, \alpha \in S(\mathfrak{F}_0)$. Take any α and f such that $\alpha = S(f)$. Since $d(\mathfrak{N}, S) = 0$ then $S(\tilde{f}(\alpha)) = S(f)$ for any $\tilde{f} \in V(f)$. Let g_0 be an arbitrary element of $\ker \mathfrak{N}(\cdot, \alpha)$. From Lemma 4.4 we know that there exists a function $h \in \mathrm{Lip}(0)$ such that $\mathfrak{N}(h(x), x) = 0$ and $h(\alpha) = g_0$.

Define $\tilde{f}(x) = f + ch(x)$ for $|c| < \delta(f)/\|h(\alpha)\|$. Then $\tilde{f} \in V(f)$ and hence $S(f + ch(\alpha)) = S(f)$. This proves that $h(\alpha) = g_0$ belongs to $G(\alpha)$ which yields $\ker \mathfrak{N}(\cdot,\alpha) \subset G(\alpha) = \ker \mathfrak{N}^*(\cdot,\alpha)$. Hence $\mathfrak{N}^* \subset \mathfrak{N}$ which completes the proof. ∎

We discuss the implications of Theorems 6.1, 6.2, and 6.3 from a computational point of view. What is the most general form of linear information operators which is permissible? Consider an idealized model in which <u>every</u> linear functional is a primitive. Then every linear information operator $\mathfrak{N}$ defined by a <u>finite</u> number of linear functionals, i.e., $\mathrm{card}(\mathfrak{N}) < +\infty$, is permissible. However even this idealized model of permissible information does not help for problems with infinite index. See also Section 8.

As we now show in several examples, the index may be infinite. Then linear information operators with finite cardinality do not supply enough information to solve the problem with an iterative algorithm.

<u>Example 6.1</u>

Assume that S is a one-to-one operator. Then $G(\alpha) = \{0\}$ and $\mathrm{codim}\, G(\alpha) = \dim \mathfrak{F}_1$. Hence

$$\mathrm{index}(S) = \dim \mathfrak{F}_1. \tag{6.8}$$

Thus if $\dim \mathfrak{F}_1 = +\infty$ then the problem S cannot be solved by an iterative algorithm using a linear information operator with finite cardinality.

Many problems can be defined by a one-to-one operator on an infinite dimensional space $\mathfrak{F}_1$. One instance is provided by the approximation problem $S(f) = f$ and $\mathfrak{F}_1 = C[a,b]$. As a second instance consider the solution of any

linear differential equation $D\alpha(x) = f_1(x)$ for $x \in \Omega$ and $\alpha(x) = f_2(x)$ for $x \in \partial\Omega$. Then $S(f) = \alpha$ where $f = (f_1, f_2)$ is a one-to-one operator. This shows that such approximation and differential equation problems cannot be solved iteratively by means of linear information operators with finite cardinality. ∎

Example 6.2

Let $\mathfrak{F}_0 = \mathfrak{F}_1 = C[a,b]$ and let k be a positive integer. Define

$$(6.9)\quad S(f) = \int_0^1 [f(t)]^k dt.$$

Since S is analytic, (6.2) yields $h \in G(\alpha)$ iff $S^{(j)}(f)h^j = 0$ for $j = 1,2,\ldots$ where $\alpha = S(f)$. Note that

$$S^{(k-1)}(f)h^{k-1} = k!\int_0^1 f(t)h^{k-1}(t)dt, \quad S^{(k)}(f)h^k = k!\int_0^1 h^k(t)dt$$

and $S^{(j)}(f) \equiv 0$ for $j > k$.

Assume that $k = 1$. Then

$$G(\alpha) = \{h: \int_0^1 h(t)dt = 0\} \text{ and } \operatorname{codim} G(\alpha) = 1.$$

Thus index(S) = 1. This means there exists a linear information operator $\mathfrak{N}$ with card($\mathfrak{N}$) = 1 such that $d(\mathfrak{N},S) = 0$. Indeed, the basic information $\mathfrak{N}^*$ defined by (6.7) is now given by $\mathfrak{N}^*(f,x) = L_1^*(f,x) = \int_0^1 f(t)dt$ and $d(\mathfrak{N}^*,S) = 0$. Furthermore the order of information $p(\mathfrak{N}^*,S) = +\infty$ and we can solve this directly. In this case, $k = 1$, the solution operator S is a linear functional and can also be used as a linear information operator whose cardinality is equal to unity. (Of course if we want to approximate S we rule out $\mathfrak{N}^* = S$ as a permissible information operator.)

Assume that $k \geq 2$. Then if k is even, $S^{(k)}(f)h^k = 0$ implies $h = 0$. If k is odd then $S^{(k-1)}(f)h^{k-1} = 0$ for a positive f implies $h = 0$. Thus $G(\alpha) = \{0\}$ and

$$\text{index}(S) = \text{codim } G(\alpha) = +\infty.$$

Hence it is impossible to iteratively approximate the value of $\int_0^1 f^k(t)dt$ by means of linear information operators with finite cardinality, for $k \geq 2$. ∎

Example 6.3 Nonlinear Equations

Let $\mathfrak{I}_1$ be the class of analytic operators, f, $f: D \subset B_1 \to B_2$ where B_1 and B_2 are real or complex Banach spaces and $N = \dim(B_1) = \dim(B_2)$. We assume that D is an open set and $\|f\| = \sup_{x \in D} \|f(x)\| < +\infty$ for $f \in \mathfrak{I}_1$. Let $\mathfrak{I}_0$ be the class of analytic operators with a unique simple zero, i.e., $f \in \mathfrak{I}_0$ iff $f \in \mathfrak{I}_1$ and there exists $\alpha \in D$ such that $f(\alpha) = 0$ and $f'(\alpha)^{-1}$ exists and is bounded. Define

(6.10) $S(f) = f^{-1}(0)$, $\mathfrak{I}_2 = B_2$.

Thus $\alpha = S(f)$ means $f(\alpha) = 0$ and the problem S is that of finding the solution of nonlinear equations. It is easy to verify that $G(\alpha) = \{h \in \mathfrak{I}_1 : h(\alpha) = 0\}$ and

$$\text{index}(S) = \dim B_1 = N.$$

Thus if $N < +\infty$ we can find a convergent linear information operator with cardinality equal to N. Indeed, the basic information $\mathfrak{N}^*$ is now given by

(6.11) $\mathfrak{N}^*(f,x) = f(x) = [f_1(x), f_2(x), \ldots, f_N(x)]$

where $f_j: D \to \mathbb{C}$ are the components of f.

For $N = +\infty$, the index of the problem S is equal to infinity. This means that we have to use linear information operators with infinite cardinality to assure the existence of iterative algorithms. For instance, $\mathfrak{N}(f,x) = f(x)$ is a convergent linear information operator. ∎

7. THE mth ORDER ITERATIONS

In the previous section we proved that if we use linear information operators with cardinality less than the index of the problem S, the class $IT(\mathfrak{N},S)$ is empty.

Let m be a real number, $m \geq 1$. We now pose and answer the following question: what is the minimal cardinality which assures the existence of a mth order iteration? See Definition 2.6. From Theorem 2.4 follows that we seek a linear information operator $\mathfrak{N}$ with minimal cardinality such that the order of information $p(\mathfrak{N},S)$ is at least equal to m. Let

$$(7.1) \quad k = k(m) = \lceil m \rceil - 1$$

In this section we assume that $S(\mathfrak{F}_0)$ is an open subset of $\mathfrak{F}_2$ and W is the class of functions $\tilde{f}$ such that (5.5) holds and the condition (5.6) is strengthened by the assumption that $\tilde{f} \in Lip(k)$.

Let H be the class of functions h such that $h: S(\mathfrak{F}_0) \to \mathfrak{F}_1$ and $h \in Lip(k)$. Define

$$(7.2) \quad G(m) = \{h: h \in H \text{ and } \forall f \in \mathfrak{F}_0, \lim_{x \to S(f)} \frac{\|S(f+ch(x))-S(f)\|}{\|x-S(f)\|^{m-\eta}} = 0,$$

$$\forall \eta > 0, \forall |c| < \delta(f)/\|h(S(f))\| \}$$

where $\delta(f)$ is defined by (5.2). Note that $S(f+ch(x))$ is well-defined for x close to $\alpha = S(f)$ since $\|ch(x)\| < \delta(f)\|h(x)\|/\|h(\alpha)\| = \delta(f)(1+o(1))$. Let $g(x) = S(f+ch(x)) - S(f)$. Then (7.2) implies that $g^{(j)}(\alpha) = 0$ for $j = 0,1,\dots,k(m)$. If S is sufficiently regular then $g^{(j)}(\alpha) = 0$ involves

conditions on $h(\alpha), h'(\alpha), \ldots, h^{(k)}(\alpha)$. Thus $G(m)$ is the set of functions h for which $S(f+ch(x))-S(f)$ has a zero α of multiplicity k.

Suppose that $A(x)$ is a linear subspace of $\mathfrak{J}_1$ such that codim $A(x) \equiv$ const for every $x \in S(\mathfrak{J}_0)$. Define

$$(7.3) \quad A = \{h\colon h \in H \text{ and } h(x) \in A(x),\ \forall\, x \in S(\mathfrak{J}_0)\}.$$

The set A contains functions from the class H whose values at x belong to $A(x)$. Clearly A is a linear subspace of H. Denote

$$(7.4) \quad \operatorname{codim} A \stackrel{\text{df}}{=} \operatorname{codim} A(x).$$

Let $A(H)$ be the class of all sets of the form (7.3). We are ready to define the mth index of S.

__Definition 7.1__

We shall say that index(S,m) is the __mth index of the problem S__ iff

$$(7.5) \quad \operatorname{index}(S,m) = \min_{A \subset G(m) \text{ and } A \in A(H)} \operatorname{codim} A \qquad \blacksquare$$

__Theorem 7.1__

Let $\mathfrak{N}\colon \mathfrak{J}_1 \times S(\mathfrak{J}_0) \to \mathfrak{J}_3$ be an arbitrary linear information operator such that $\mathfrak{N} \in \operatorname{Lip}(k(m))$ and $\operatorname{card}(\mathfrak{N}) < \operatorname{index}(S,m)$. Then

$$(7.6) \quad p(\mathfrak{N},S) < m. \qquad \blacksquare$$

__Proof__

Assume that there exists an information operator $\mathfrak{N}$ such that $\mathfrak{N} \in \operatorname{Lip}(k(m))$, $n = \operatorname{card}(\mathfrak{N}) < \operatorname{index}(S,m)$ and $p(\mathfrak{N},S) \geq m$. Let $\mathfrak{N} \simeq [L_1, L_2, \ldots, L_n]$ where $L_i \in \operatorname{Lip}(k)$, $i = 1,2,\ldots,n$. Define

(7.7) $K = \{h\colon h \in H \text{ and } L_i(h(x),x) = 0,\ i = 1,2,\ldots,n,\ \forall\, x \in S(\mathfrak{J}_0)\}$.

Since $L_1(\cdot,x),\ldots,L_n(\cdot,x)$ are linearly independent for every $x \in S(\mathfrak{J}_0)$ then $K \in A(H)$ and codim $K = n$. We show that $K \subset G(m)$. Let h be an arbitrary element of K. Take any $f \in \mathfrak{J}_0$, $\alpha = S(f)$, and define $\tilde{f}(x) = f+ch(x)$ for $|c| < \delta(f)/\|h(\alpha)\|$. Then $\tilde{f}(x) \in \mathfrak{J}_0$ for x close to α and $\mathfrak{N}(\tilde{f}(x),x) = \mathfrak{N}(f,x)$ since $h(x) \in \ker \mathfrak{N}(\cdot,x)$. Thus $\tilde{f} \in V(f)$. Since $p(\mathfrak{N},S) \geq m$, Definition 2.5 yields $S(f+ch(x)) - S(f) = o(\|x-\alpha\|^{m-\eta})$ for every $\eta > 0$. This means that $h \in G(m)$ and implies $K \subset G(m)$. From (7.5) we get

$$\text{index}(S,m) \leq \text{codim } K = n = \text{card}(\mathfrak{N}).$$

This is a contradiction. Hence $p(\mathfrak{N},S) < m$ which completes the proof. ■

Theorem 7.1 states that any regular linear information operator $\mathfrak{N}$ with order of information $p(\mathfrak{N},S)$ greater or equal to m has to have cardinality at least equal to the mth index of the problem S.

We now construct a linear information operator $\mathfrak{N}$, $\mathfrak{N} \in \text{Lip}(k)$ and $\text{card}(\mathfrak{N}) = \text{index}(S,m)$ such that $p(\mathfrak{N},S) \geq m$.

Let $A^* = \{h\colon h \in H \text{ and } h(x) \in A^*(x),\ \forall\, x \in S(\mathfrak{J}_0)\}$ be a linear subspace of H such that the minimum in (7.5) is attained for A^*, i.e.,

(7.8) $A^* \subset G(m)$ and codim $A^* = \text{index}(S,m)$.

This means that codim $A^*(x) = \text{index}(S,m)$ for every $x \in S(\mathfrak{J}_0)$. Decompose $\mathfrak{J}_1 = A^*(x) \oplus A^*(x)$, i.e., for every $f \in \mathfrak{J}_1$

(7.9) $$f = f_0 + \sum_{i=1}^{n^*} L_i^*(f,x)\xi_i(x)$$

where $n^* = \text{index}(S,m)$, $f_0 \in A^*(x)$ and $L_1^*(\cdot,x),\ldots, L_{n^*}^*(\cdot,x)$ are linearly independent linear functionals for every $x \in S(\mathfrak{I}_0)$.

Let

$$(7.10)\quad \mathfrak{N}_m^*(f,x) = [L_1^*(f,x),\ldots,L_{n^*}^*(f,x)]$$

be an <u>mth basic linear information operator</u>.

<u>Definition 7.2</u>

We shall say a <u>solution operator S belongs to the class Lip(k(m))</u>, $S \in \text{Lip}(k(m))$, iff there exists $\mathfrak{N}_m^*$ of the form (7.10) such that $\mathfrak{N}_m^* \in \text{Lip}(k(m))$. ■

Note that $\mathfrak{N}^* \in \text{Lip}(k)$ means that $L_i^* \in \text{Lip}(k)$ for $i = 1,2,\ldots,n^*$. Clearly $\text{card}(\mathfrak{N}_m^*) = \text{index}(S,m)$. From (7.9) we have $L_i^*(\xi_j(x),x) = \delta_{ij}$ and we can assume that the functions $\xi_j \in \text{Lip}(k)$.

We estimate the order $p(\mathfrak{N}_m^*,S)$ of the mth basic linear information $\mathfrak{N}_m^*$.

<u>Theorem 7.2</u>

Assume that $S \in \text{Lip}(k(m))$. Then

$$(7.11)\quad p(\mathfrak{N}_m^*,S) \geq m. \qquad ■$$

<u>Proof</u>

Take any $f \in \mathfrak{I}_0$, $\alpha = S(f)$, and any $\tilde{f} \in V(f)$. Then $h(x) = \tilde{f}(x) - f$ belongs to $\ker \mathfrak{N}_m^*(\cdot,x)$, i.e., $L_i^*(h(x),x) = 0$. Since $\tilde{f} \in \text{Lip}(k)$ then also $h \in \text{Lip}(k)$. From this and (7.9) we get that $h(x) \in A^*(x)$ for every $x \in S(\mathfrak{I}_0)$. This means $h \in A^*$. Since $A^* \subset G(m)$, $h \in G(m)$. Note that $\|h(\alpha)\| < \delta(f)$ which implies that we can put $c = 1$ in (7.2) getting

$$\|S(\tilde{f}(x))-S(f)\| = o(\|x-\alpha\|^{m-\eta}),\quad \forall\, \eta > 0.$$

This proves that $p(\mathfrak{N}_m^*,S) \geq m$. ■

Theorems 7.1 and 7.2 state that we have to use linear information operators with cardinality at least equal to the mth index to assure the existence of a mth order iteration. If $\text{index}(S,m) = +\infty$ then linear information operators with finite cardinality do not supply enough information to guarantee the existence of a mth order iteration.

As in Section 6 we prove that in some cases any linear information operator $\mathfrak{N}$ with order of information at least equal to m contains the information $\mathfrak{N}_m^*$.

<u>Theorem 7.3</u>

Let $G(m)$ be a linear set such that $G(m) = A^*$. Let $\mathfrak{N}$ be a linear information operator such that $\mathfrak{N} \in \text{Lip}(k(m))$. Then $p(\mathfrak{N},S) \geq m$ implies $\mathfrak{N}_m^* \subset \mathfrak{N}$. ■

<u>Proof</u>

Take any $f \in \mathfrak{J}_0$, $\alpha = S(f)$. Since $p = p(\mathfrak{N},S) \geq m$ then $S(\tilde{f}(x)) - S(f) = o(\|x-\alpha\|^{p-\eta})$, $\forall\, \eta > 0$, for any $\tilde{f} \in V(f)$. Let g_0 be an arbitrary element of $\ker \mathfrak{N}(\cdot,\alpha)$. From Lemma 4.4 we can find a function h such that $h \in \text{Lip}(k)$, $\mathfrak{N}(h(x),x) \equiv 0$ and $h(\alpha) = g_0$. Since $S(f+ch(x)) - S(f) = o(\|x-\alpha\|^{m-\eta})$, $\forall\, \eta > 0$, and $|c| < \delta(f)/\|h(\alpha)\|$ then $h \in G(m) = A^*$. Thus $h(x) \in A^*(x)$, i.e., $L_i^*(h(\alpha),\alpha) \equiv 0$. This implies that $g_0 = h(\alpha) \in \ker \mathfrak{N}_m^*(\cdot,\alpha)$ and $\ker \mathfrak{N}(\cdot,\alpha) \subset \ker \mathfrak{N}_m^*(\cdot,\alpha)$. Since α is arbitrary, $\mathfrak{N}_m^* \subset \mathfrak{N}$ which completes the proof. ■

Compare with Theorem 4.2 in Traub and Woźniakowski [76b] where a similar problem is considered for nonlinear equations.

We illustrate the concept of the mth index by two examples.

<u>Example 7.1</u> Nonlinear Equations

As in Example 6.3 define $S(f) = f^{-1}(0)$. Let $h \in H$. Now $h(x)$ is a function from $\mathfrak{J}_0$, i.e., $h(x) = h(x,\cdot)$. It is easy to verify that

$$G(m) = \{h: \frac{\partial^j h(\alpha,\alpha)}{\partial x^j} = 0, \quad j = 0,1,\ldots,k(m), \quad \forall\, \alpha \in S(\mathfrak{J}_0)\}$$

where $k = k(m)$ is defined by (7.1).

This means that $h(x,t) = O((x-t)^{k+1})$ belongs to $G(m)$. The mth index is given by

$$\text{index}(S,m) = N\binom{N+k}{k}$$

where N, $N < +\infty$, is the dimension of the problem. For $N = +\infty$, $\text{index}(S,m) = +\infty$. The information $\mathfrak{N}_m^*$ is now given by

$$\mathfrak{N}_m^*(f,x) = \{f(x),f'(x),\ldots,f^{(k)}(x)\}$$

and was intensively studied by Traub and Woźniakowski [76b, 77a, 77b]. ■

<u>Example 7.2</u> Linear Equations

Consider, as in Example 7.1, $S(f) = f^{-1}(0)$ with the additional assumption that f is an affine function, i.e., $f(t) = At - b$ where A is a nonsingular $N\times N$ matrix and b is a $N\times 1$ vector. Then

$$G(m) = \{h: \frac{\partial^j h(\alpha,\alpha)}{\partial x^j} = 0, \quad j = 0,1,\ldots,k(m), \quad \forall\, \alpha \in S(\mathfrak{J}_0)\}$$

where $h(x,t)$ is a linear function of t. It is easy to check that

$$\text{index}(S,m) = N(N+1), \quad \forall\, m > 1.$$

From this follows that every algorithm φ which uses any linear information operator $\mathfrak{N}$ with cardinality less than $N(N+1)$ has order $p(\varphi) \le p(\mathfrak{N},S) \le 1$. Furthermore every interpolatory algorithm φ^I which uses the linear information operator $\mathfrak{N}^*(f,x) = [f(x),f'(x)]$, $\mathrm{card}(\mathfrak{N}^*) = N(N+1)$, has order $p(\varphi^I) = p(\mathfrak{N}^*,S) = +\infty$. Since $f(x) = Ax - b$ and $f'(x) \equiv A$, f is fully determined by $\mathfrak{N}^*(f,x)$ and φ^I requires the solution of a linear system and φ^I is a direct algorithm. ∎

8. COMPLEXITY INDEX FOR ITERATIVE LINEAR INFORMATION

We specify our model of computation for iterative linear information operators which is similar to the model for the linear case in Part A.

<u>Model of Computation for Iterative Linear Information</u>

(i) Let P be a given collection of primitives. We assume that the addition of two elements of $\mathfrak{F}_2$, f+g, and the multiplication of an element of $\mathfrak{F}_2$ by a scalar, cf, are primitive operations which belong to P. We also assume that every linear functional $L(\cdot,x)$ is a primitive operation which belongs to P for every x under consideration. This implies that any linear information operator $\mathfrak{N} = [L_1,L_2,\ldots,L_n]$ of finite cardinality is <u>permissible</u> where $L_1,L_2,\ldots,L_n$ are arbitrary linear functionals.

(ii) To normalize the complexity measure we assume that the cost of the addition of two elements of $\mathfrak{F}_2$ and the multiplication of an element of $\mathfrak{F}_2$ by a scalar is taken as unity. Assume that the complexity of evaluating a linear functional $L(\cdot,x)$ does not depend on x and let comp(L) = comp(L(f,x)). Let $\mathfrak{N} = [L_1,\ldots,L_n]$ be a linear information operator, card($\mathfrak{N}$) = n. We assume that $\mathfrak{N}(f,x)$ is computed by the independent evaluation of $L_1(f,x),\ldots,L_n(f,x)$ and the information complexity of $\mathfrak{N}$ is given by

$$(8.1)\quad \mathrm{comp}(\mathfrak{N}) = \sum_{i=1}^{n} \mathrm{comp}(L_i).$$

If $\mathrm{comp}(L_i) \equiv c_1$ then $\mathrm{comp}(\mathfrak{N}) = nc_1$ which shows how the information complexity depends on the cardinality of $\mathfrak{N}$.

(iii) Let φ be a permissible algorithm which uses $\mathfrak{N}$ and finds an ε'-approximation to $\alpha = S(f)$. Let comb(φ) be the combinatory complexity of φ. For all problems of practical interest, φ has to use every $L_i(f,x)$, $i = 1,2,\ldots,n$, and the current approximation x at least once and therefore comb(φ) $\geq n$. We rule out special problems and information operators, assuming that comb(φ) $\geq n$ for every algorithm under consideration. ∎

We analyze the complexity index of iterative algorithms using a linear information operator $\mathfrak{N}$. Assume that $p(\mathfrak{N},S)$ is greater than unity where $\mathfrak{N} = [L_1,\ldots,L_n]$. Let φ be a permissible algorithm from the class $\Phi_{perm}(\mathfrak{N},S)$ and $p(\varphi) > 1$. See Section 3. Therefore φ also belongs to the class $IT(\mathfrak{N},S)$ of iterative algorithms. See Corollary 5.1. The complexity index $z(\varphi)$ is defined (compare with (3.11)) as

$$(8.2)\quad z(\varphi) = \frac{\text{comp}(\mathfrak{N}) + \text{comb}(\varphi)}{\log p(\varphi)}$$

where $\text{comp}(\mathfrak{N}) = \sum_{i=1}^{n} \text{comp}(L_i)$ is the information complexity and

$$\text{comb}(\varphi) = \sup_{(f,x)} \text{comp}(\varphi(x,\mathfrak{N}(f,x))$$

is the combinatory complexity of φ. For simplicity assume that $\text{comp}(L_i) \equiv c_1$. From Theorem 2.3 we know that $p(\varphi) \leq p(\mathfrak{N},S)$. Since comb($\varphi$) $\geq n$, this yields

$$(8.3)\quad z(\varphi) \geq \frac{nc_1 + n}{\log p(\mathfrak{N},S)} .$$

Furthermore, Theorem 2.4 guarantees the existence of algorithms whose order is equal to the order of information. Assume that one of these maximal order algorithms is permissible. Let $\overline{\text{comb}}(\mathfrak{N},S)$ denote the minimal combinatory complexity of algorithms with maximal order, i.e.,

$$(8.4)\quad \overline{\text{comb}}(\mathfrak{N},S) = \inf_{\varphi:\ p(\varphi)=p(\mathfrak{N},S)} \text{comb}(\varphi).$$

Let $\underline{\text{comb}}(\mathfrak{N},S)$ denote the minimal combinatory complexity of algorithms φ from $\Phi_{perm}(\mathfrak{N},S)$ with $p(\varphi) > 1$, i.e.,

$$(8.5)\quad \underline{\text{comb}}(\mathfrak{N},S) = \inf_{\varphi:\ p(\varphi)>1} \text{comb}(\varphi).$$

Of course $n \le \underline{\text{comb}}(\mathfrak{N},S) \le \overline{\text{comb}}(\mathfrak{N},S)$. From (8.4) and (8.5) we get bounds on the complexity index $z(\mathfrak{N},S)$ *of the information* $\mathfrak{N}$ *for the problem* S. See Definition 3.1.

<u>Lemma 8.1</u>

$$(8.6)\quad \frac{nc_1 + \underline{\text{comb}}(\mathfrak{N},S)}{\log p(\mathfrak{N},S)} \le z(\mathfrak{N},S) \le \frac{nc_1 + \overline{\text{comb}}(\mathfrak{N},S)}{\log p(\mathfrak{N},S)}\ . \qquad \blacksquare$$

Note that if $nc_1 \gg \overline{\text{comb}}(\mathfrak{N},S)$ then

$$(8.7)\quad z(\mathfrak{N},S) \cong \frac{nc_1}{\log p(\mathfrak{N},S)}$$

and every maximal order algorithm φ is close to a minimal complexity index algorithm since $z(\varphi) \cong nc_1/\log p(\mathfrak{N},S)$.

The above discussion motivates the following problem. For fixed n find a linear information operator $\mathfrak{N}$ with $\text{card}(\mathfrak{N}) \le n$ and maximal order of information. Let Ψ_n be the class of all linear information operators $\mathfrak{N}$ which are sufficiently regular and $\text{card}(\mathfrak{N}) \le n$. We assume that S is also sufficiently regular.

<u>Definition 8.1</u>

We shall say $p(n,S)$ is the <u>nth maximal order for the problem S</u> iff

(8.8) $p(n,S) = \sup_{\mathfrak{N}\in\Psi_n} p(\mathfrak{N},S)$.

We shall say $\mathfrak{N}^{mo}, \mathfrak{N}^{mo} \in \Psi_n$, is a <u>nth maximal order information for the problem S</u> iff

(8.9) $p(\mathfrak{N}^{mo},S) = p(n,S)$. ■

The results of Sections 6 and 7 enable us to find the nth maximal order and the nth maximal order information. Recall that index(S,m) is the mth index of the problem S, $m \geq 1$. See Definition 7. Note that index(S,m) is a nondecreasing function of m.

Suppose first that index(S,1) $= +\infty$. Then index(S,m) $\equiv +\infty$ and Theorem 7.1 yields $p(\mathfrak{N},S) = 0$ for every $\mathfrak{N}$ with card($\mathfrak{N}$) $< +\infty$. Thus $p(n,S) \equiv 0$ and every linear information operator satisfies (8.9).

Suppose then that

(8.10) $n_0 = \text{index}(S,1) < +\infty$.

From Theorem 7.1 we immediately get $p(n,S) = 0$ for $n < n_0$. For $n \geq n_0$ define

(8.11) $q(n,S) = \inf\{m: n < \text{index}(S,m)\}$

with the convention $q(n,S) = +\infty$ if index(S,m) $\leq n$, $\forall$ m. The function $q(n,S)$ is a nondecreasing function of n and $q(n_0,S) \geq 1$. We are ready to prove

Theorem 8.1

Let n_0 = index(S,1). Then

$$(8.12)\quad p(n,S) = \begin{cases} 0 & \text{for } n < n_0, \\ q(n,S) & \text{for } n \geq n_0. \end{cases}$$ ■

Proof

It suffices to prove (8.12) for $n_0 < +\infty$ and $n \geq n_0$. Let $q = q(n,S)$. Take any positive η. From (8.11) index$(S,q-\eta) \leq n$. Theorem 7.2 guarantees that there exists a linear information operator $\mathfrak{N}$ such that card$(\mathfrak{N})$ = index$(S,q-\eta) \leq n$ and $p(\mathfrak{N},S) \geq q-\eta$. Since η is arbitrary, $p(n,S) \geq q = q(n,S)$.

From (8.11) we have $n <$ index$(S,q+\eta)$. Theorem 7.1 states that for any $\mathfrak{N} \in \Psi_n$, $p(\mathfrak{N},S) < q+\eta$. Thus $p(n,S) \leq q$. Hence $p(n,S) = q(n,S)$ which completes the proof. ■

Theorem 8.1 states the dependence of the nth maximal order $p(n,S)$ on the problem S. Since we assume S and the linear information operators $\mathfrak{N}$ are sufficiently regular, the nth maximal order $p(n,S)$ is an integer. We find a nth maximal order information for $n \geq n_0$ = index(S,1) where $n_0 < +\infty$. Let $\eta \in (0,1)$ and let $\mathfrak{N}_n$ be a linear information operator such that card$(\mathfrak{N}_n)$ = index$(S,p(n,S)-\eta) \leq n$ and $p(\mathfrak{N}_n,S) \geq p(n,S)-\eta > p(n,S) - 1$. The existence of $\mathfrak{N}_n$ is guaranteed by Theorem 7.2. Since $p(\mathfrak{N}_n,S)$ is an integer, we get $p(\mathfrak{N}_n,S) = p(n,S)$. This proves the following theorem.

Theorem 8.2

Let n_0 = index$(S,1) < +\infty$. For $n \geq n_0$, $\mathfrak{N}_n$ is a nth maximal order information for the problem S. ■

We examine the complexity index $z(\mathfrak{N},S)$ of the information operators $\mathfrak{N}$ from the class Ψ_n.

Definition 8.2

We shall say $z(n,S)$ is the nth minimal complexity index for the problem S iff

$$(8.13)\quad z(n,S) = \inf_{\mathfrak{N}\in\Psi_n} z(\mathfrak{N},S).$$

We shall say that $n_{opt} = n_{opt}(S)$ is the optimal cardinality number for the problem S with respect to the complexity index iff

$$(8.14)\quad z(n_{opt},S) = \inf_n z(n,S).$$

We shall say a linear information operator $\mathfrak{N}^{oi}$ is an optimal information operator for the problem S iff

$$(8.15)\quad \mathrm{card}(\mathfrak{N}^{oi}) = n_{opt}(S) \text{ and } p(\mathfrak{N}^{oi},S) = p(n_{opt},S).$$ ■

Let $\underline{\mathrm{comb}}(n,S) = \inf_{\mathfrak{N}\in\Psi_n} \underline{\mathrm{comb}}(\mathfrak{N},S)$ and $\overline{\mathrm{comb}}(n,S) = \inf_{\mathfrak{N}\in\Psi'_n} \overline{\mathrm{comb}}(\mathfrak{N},S)$ (where $\Psi'_n = \{\mathfrak{N}: \mathfrak{N} \in \Psi_n \text{ and } p(\mathfrak{N},S) = p(n,S)\}$) be the minimal combinatory complexity of $\underline{\mathrm{comb}}(\mathfrak{N},S)$ and the minimal combinatory complexity of $\overline{\mathrm{comb}}(\mathfrak{N},S)$ for the information operators with the maximal order $p(n,S)$ respectively. See (8.4) and (8.5). From Lemma 8.1 and Theorems 8.1 and 8.2 we get the following estimates of the nth minimal complexity index.

Lemma 8.2

$$(8.16)\quad \frac{nc_1 + \underline{\mathrm{comb}}(n,s)}{\log p(n,S)} \le z(n,S) \le \frac{nc_1 + \overline{\mathrm{comb}}(n,S)}{\log p(n,S)}.$$ ■

Observe that (8.16) is simpler if $nc_1 >> \overline{\text{comb}}(n,s)$. Then

$$(8.17)\quad z(n,S) \cong \frac{nc_1}{\log p(n,S)} .$$

This shows that the nth minimal complexity index depends on the cardinality of information and the nth maximal order. To find the optimal cardinality number one seeks the minimum of the function $f(n) = z(n,S)$. The value of $n_{opt}(S)$ depends only on how fast the functions $\underline{\text{comb}}(n,S)$, $\overline{\text{comb}}(n,S)$ and $p(n,S)$ tend to infinity with n. Knowing $n_{opt}(S)$ one can easily find an optimal information operator $\mathfrak{N}^{oi}$ from Theorem 8.2. We conclude this section by an example.

Example 8.1

Consider the solution of nonlinear equations, $S = f^{-1}(0)$ where N is the dimension of the problem, $N < +\infty$. Traub and Woźniakowski [77b] showed that

$$n_{opt}(S) = \begin{cases} 3 & \text{for } N = 1, \text{ the scalar case,} \\ N(N+1) & \text{for } N \geq 2, \text{ the multivariate case.} \end{cases}$$

The optimal information operator $\mathfrak{N}^{oi}(f,x) = [f(x),f'(x),f''(x)]$ for $N = 1$ and $\mathfrak{N}^{oi}(f,x) = [f(x),f'(x)]$ for $N \geq 2$. ∎

CHAPTER III

EXTENSIONS AND COMMENTS

9. INFORMATION OPERATOR WITH MEMORY

In this section we briefly indicate how the concepts of the limiting diameter and order of information can be generalized for information operators with memory. As always we want to approximate the solution $\alpha = S(f)$. Suppose that $x_0, x_1, \ldots, x_r$, $r \geq 1$, are known distinct approximations to α. Let

$$(9.1) \quad \mathfrak{N}: D_{\mathfrak{N}} \subset \mathfrak{J}_1 \times \mathfrak{J}_2^{r+1} \to \mathfrak{J}_3$$

be an information operator with memory (not necessarily linear) where $(f, x_0, x_1, \ldots, x_r) \in D_{\mathfrak{N}}$ for all $f \in \mathfrak{J}_0$ and all distinct $x_0, x_1, \ldots, x_r$ sufficiently close to α, and $\mathfrak{J}_3$ is a given space. The parameter r measures the size of the memory used by $\mathfrak{N}$. In general $\mathfrak{N}(f, x_0, x_1, \ldots, x_r)$ does not uniquely define the solution α and many different problem elements have the same information at $x_0, x_1, \ldots, x_r$ as f. We define the concept of equality with respect to $\mathfrak{N}$.

<u>Definition 9.1</u>

We shall say <u>$\tilde{f} = \tilde{f}(x_0, x_1, \ldots, x_r)$ is equal to f with respect to $\mathfrak{N}$</u> iff

(i) $\tilde{f}: D_{\tilde{f}} \subset \mathfrak{J}_2^{r+1} \to \mathfrak{J}_0, \quad \tilde{f} \in W$

where W is a given space and there exists $\Gamma = \Gamma(f) > 0$ such that $[J(\Gamma)]^{r+1} \subset D_{\tilde{f}}$ where $J(\Gamma) = \{x: \|x - \alpha\| \leq \Gamma\}$ with $\alpha = S(f)$,

(ii) $\mathfrak{N}(\tilde{f}(x_0, x_1, \ldots, x_r), x_0, x_1, \ldots, x_r) = \mathfrak{N}(f, x_0, x_1, \ldots, x_r), \ \forall (x_0, x_1, \ldots, x_r) \in D_{\tilde{f}}$.

For brevity we write $\tilde{f} \in V(f)$ where $V(f)$ is the set of all functions $\tilde{f}$ which are equal to f with respect to $\mathfrak{N}$. ∎

The space W describes the regularity of $\tilde{f}$ with respect to $x_0, x_1, \ldots, x_r$. The limiting diameter $d(\mathfrak{N},S)$ of $\mathfrak{N}$ for the problem S is defined as follows.

Definition 9.2

We shall say $d(\mathfrak{N},S)$ is the limiting diameter of information $\mathfrak{N}$ for the problem S iff

$$(9.2)\quad d(\mathfrak{N},S) = \sup_{f\in\mathfrak{F}_0} \ \sup_{\tilde{f}_1,\tilde{f}_2\in V(f)} \ \limsup_{\substack{x_i\to\alpha \\ i=0,1,\ldots,r}} \| S(\tilde{f}_1(x_0,\ldots,x_r)) - S(\tilde{f}_2(x_0,\ldots,x_r))\|.$$

We shall say the information $\mathfrak{N}$ is convergent for the problem S (or simply $\mathfrak{N}$ is convergent) iff

$$(9.3)\quad d(\mathfrak{N},S) = 0. \qquad \blacksquare$$

Let

$$(9.4)\quad \varphi: D_\varphi \subset \mathfrak{F}_2^{r+1} \times \mathfrak{N}(D_{\mathfrak{N}}) \to \mathfrak{F}_2$$

be a stationary algorithm with memory which generates the sequence of approximations by

$$(9.5)\quad x_{j+1} = \varphi(x_j, x_{j-1}, \ldots, x_{j-r}, \quad \mathfrak{N}(f, x_j, x_{j-1}, \ldots, x_{j-r}))$$

for $j = r, r+1, \ldots$. Note that φ uses r previously computed approximations and the information $\mathfrak{N}$ computed at them. Usually $\mathfrak{N}$ consists of "new information" at x_j and reuses previously computed information at $x_{j-1}, \ldots, x_{j-r}$. Let $\Phi(\mathfrak{N},S)$ be the class of all stationary algorithms with memory. We want to examine the convergence of the sequence $\{x_j\}$ to α. As in Section 2 we define the limiting error of φ as follows.

Definition 9.3

We shall say $e(\varphi)$ is the limiting error of algorithm φ iff

(9.6) $$e(\varphi) = \sup_{f \in \mathfrak{J}_0} \sup_{\tilde{f} \in V(f)} \limsup_{\substack{x_i \to \alpha \\ i=0,1,\dots,r}} \| \varphi(x_0,\dots,x_r,\mathfrak{N}(f,x_0,\dots,x_r)) - S(\tilde{f}(x_0,\dots,x_r)) \|.$$

The algorithm φ is called convergent iff $e(\varphi) = 0$. ■

Using the same argument as in Theorem 2.1 it is easy to prove

Theorem 9.1

For any algorithm φ, $\varphi \in \Phi(\mathfrak{N},S)$,

(9.7) $$e(\varphi) \geq \frac{1}{2}\, d(\mathfrak{N},S).$$ ■

Thus, $\frac{1}{2}\, d(\mathfrak{N},S)$ is the inherent error of information $\mathfrak{N}$ for any algorithm φ. The diameter $d(\mathfrak{N},S)$ is an upper bound on "interpolatory algorithms" which are defined analogously as in Section 2.

Definition 9.4

We shall say φ^I, $\varphi^I \in \Phi(\mathfrak{N},S)$, is an interpolatory algorithm iff

(9.8) $$\varphi(x_0,\dots,x_r, \mathfrak{N}(f,x_0,\dots,x_r)) = S(\tilde{f}(x_0,\dots,x_r))$$

for some $\tilde{f} \in V(f)$. ■

This means that an interpolatory algorithm φ constructs as the new approximation to α the exact solution of a problem element $\tilde{f}(x_0,\dots,x_r)$ which shares the same information as f at $x_0,\dots,x_r$. In practice $\tilde{f}(x_0,\dots,x_r)$ should be "simpler" than f. It is straightforward to prove

Theorem 9.2

For any interpolatory algorithm φ^I, $\varphi^I \in \Phi(\mathfrak{N},S)$,

$$e(\varphi^I) \le d(\mathfrak{N},S). \tag{9.9}$$ ■

Thus, there exists a convergent algorithm in $\Phi(\mathfrak{N},S)$ iff $\mathfrak{N}$ is convergent. Assume then that $\mathfrak{N}$ is a convergent information operator. We examine how fast $S(\tilde{f}(x_0,\ldots,x_r))$ tends to $S(f)$ as x_i approaches α for $i = 0,1,\ldots,r$. We generalize the concept of order of information as follows. Let A be a set of real $(r+1)$-tuples defined by

$$A = \{(q_0,q_1,\ldots,q_r):\ q_i \ge 0,\ \sum_{i=0}^{r} q_i \ge 1 \text{ such that} \tag{9.10}$$

$$\forall f \in \mathfrak{F}_0,\ \alpha = S(f), \text{ and } \forall \tilde{f} \in V(f) \text{ we have}$$

$$\lim_{\|x_0-\alpha\| \le \ldots \le \|x_r-\alpha\| \to 0} \frac{\|S(\tilde{f}(x_0,\ldots,x_r))-S(f)\|}{\|x_0-\alpha\|^{q_0-\eta} \cdot \ldots \cdot \|x_r-\alpha\|^{q_r-\eta}} = 0,\ \forall \eta > 0\}.$$

Suppose for a moment that A is non-empty and define

$$t(A) = \sup\{t:\ t^{r+1} = q_0 t^r + q_1 t^{r-1} + \ldots + q_r \text{ for } (q_0,q_1,\ldots,q_r) \in A\}. \tag{9.11}$$

Note that the polynomial $t^{r+1} - (q_0 t^r + \ldots + q_r)$ for $q_i \ge 0$ and $\sum_{i=0}^{r} q_i \ge 1$ has a unique positive zero $t_1 \ge 1$. Therefore $t(A) \ge 1$.

Definition 9.5

We shall say $p(\mathfrak{N},S)$ is the <u>order of information $\mathfrak{N}$ for the problem S</u> iff

$$p(\mathfrak{N},S) = \begin{cases} 0 & \text{if A is empty,} \\ t(A) & \text{otherwise.} \end{cases} \tag{9.12}$$ ■

For $r = 0$, i.e., for information operators without memory, the set A coincides with the set defined by (2.17) and $t(A) = \sup\{q: q \in A\}$. Therefore Definition 9.5 for $r = 0$ coincides with Definition 2.5. See also Kung and Traub [76] where a similar technique for measuring the speed of convergence is proposed.

We state a theorem which often helps the calculation of the order of information $p(\mathfrak{N},S)$.

Theorem 9.3

If there exist $p_0, p_1, \ldots, p_r$ such that $p_i \geq 0$, $\sum_{i=0}^{r} p_i \geq 1$, and

(i) for all $f \in \mathfrak{I}_0$ and all $\tilde{f} \in V(f)$,

$$(9.13) \quad \| S(\tilde{f}(x_0,\ldots,x_r)) - S(f) \| \leq c_1(f,\tilde{f}) \| x_0-\alpha\|^{p_0} \cdot \ldots \cdot \| x_r - \alpha\|^{p_r}$$

for all $x_0,\ldots,x_r$ sufficiently close to $\alpha = S(f)$, $\| x_0-\alpha\| \leq \ldots \leq \| x_r-\alpha\|$ and $c_1(\tilde{f},f) < +\infty$, then

$$(9.14) \quad p(\mathfrak{N},S) \geq p$$

where p is the unique positive zero of $t^{r+1} - (p_0t^r + \ldots + p_r)$.

(ii) If additionally there exist $f \in \mathfrak{I}_0$ and $\tilde{f} \in V(f)$ such that

$$(9.15) \quad \| S(\tilde{f}(x_0,\ldots,x_r)) - S(f) \| \geq c_2(f,\tilde{f}) \| x_0-\alpha\|^{p_0} \cdot \ldots \cdot \| x_r - \alpha\|^{p_r}$$

for all $x_0,\ldots,x_r$ sufficiently close to $\alpha = S(f)$, $\| x_0-\alpha\| \leq \ldots \leq \| x_r-\alpha\|$ and $c_2(\tilde{f},f) > 0$, then

$$(9.16) \quad p(\mathfrak{N},S) = p. \qquad \blacksquare$$

Proof

From (9.13) it follows that $(p_0, p_1, \ldots, p_r) \in A$. Thus $p(\mathfrak{N}, S) = t(A) \geq p$ due to (9.11). This proves (9.14).

Suppose that (9.15) holds. Let $(q_0, q_1, \ldots, q_r) \in A$. Then

$$(9.17)\quad c_2(f,\tilde{f})\, \|x_0-\alpha\|^{p_0} \cdot \ldots \cdot \|x_r-\alpha\|^{p_r} \leq \|S(\tilde{f}(x_0,\ldots,x_r))-S(f)\| =$$

$$o(\|x_0-\alpha\|^{q_0-\eta} \cdot \ldots \cdot \|x_r-\alpha\|^{q_r-\eta}),\ \forall\, \eta > 0.$$

We need the following lemma.

Lemma 9.1

Let $\omega(t) = p_0 t^r + \ldots + p_r$, $u(t) = q_0 t^r + \ldots + q_r$ where $p_i \geq 0$, $q_i \geq 0$ and $\sum_{i=0}^{r} q_i \geq 1$. Let p and q be the unique positive zeros of $t^{r+1} - \omega(t)$ and $t^{r+1} - u(t)$ respectively. Then

$$(9.18)\quad p_0 + p_1 + \ldots + p_i \geq q_0 + q_1 + \ldots + q_i,\ \forall\, i \in [0,r],$$

implies $p \geq q$. ■

Proof

Let $\omega(t) = \omega(1) + \omega'(1)(t-1) + \ldots + \frac{1}{r!}\omega^{(r)}(1)(t-1)^r$. Note that $\omega^{(k)}(1) = \sum_{i=0}^{r-k} p_i c_i$ where $c_i = (r-i)!/(r-i-k)!$ for $i = 0,1,\ldots,r-k$. Set $c_{r-k+1} = 0$. It is easy to see that $c_i > c_{i+1}$ for $i = 0,1,\ldots,r-k$. Multiply the inequality $p_0 + p_1 + \ldots + p_j \geq q_0 + q_1 + \ldots + q_j$ by $(c_j - c_{j+1})$ and add for $j = 0,1,\ldots,r-k$. Then

$$\omega^{(k)}(1) = \sum_{i=0}^{r-k} p_i c_i = \sum_{i=0}^{r-k} p_i \sum_{j=i}^{r-k} (c_j - c_{j+1}) = \sum_{j=0}^{r-k} (c_j - c_{j+1}) \sum_{i=0}^{j} p_i \geq$$

$$\geq \sum_{j=0}^{r-k} (c_j - c_{j+1}) \sum_{i=0}^{j} q_i = u^{(k)}(1)$$

for $k=0,1,\ldots,r-1$. This implies $\omega(t) \geq u(t)$ for $t \geq 1$. Set $t = q$. Then $\omega(q) \geq u(q) = q^r$. Thus $q^r - \omega(q) \leq 0$ which yields $p \geq q$. ∎

Let i be any integer from $[0,r]$. Assume that $x_{i+1},\ldots,x_r$ are fixed and let $x_1 = x_2 = \ldots = x_i$ tend to α. Then (9.17) yields $p_0 + p_1 + \ldots + p_i \geq q_0 + q_1 + \ldots + q_i$. From Lemma 9.1 and (9.11) we get $p(\mathfrak{N},S) \leq \sup q \leq p$. Due to (9.14), $p(\mathfrak{N},S) = p$ which proves (9.16). ∎

Let φ be a convergent algorithm from $\Phi(\mathfrak{N},S)$. The order of φ is defined in a similar way as the order of information. Let B be a set of real $(r+1)$-tuple sequences defined by

$$(9.19)\quad B = \Big\{(q_0,q_1,\ldots,q_r):\ q_i \geq 0;\ \sum_{i=0}^{r} q_i \geq 1 \text{ such that}$$

$$\forall f \in \mathfrak{F}_0,\ \alpha = S(f), \text{ and } \forall \tilde{f} \in V(f) \text{ we have}$$

$$\lim_{\|x_0-\alpha\| \leq \ldots \leq \|x_r-\alpha\| \to 0} \frac{\|\varphi(x_0,\ldots,x_r,\mathfrak{N}(f,x_0,\ldots,x_r)) - S(\tilde{f}(x_0,\ldots,x_r))\|}{\|x_0-\alpha\|^{q_0-\eta} \cdot \ldots \cdot \|x_r-\alpha\|^{q_r-\eta}} = 0$$

$$\forall \eta > 0\Big\}.$$

Suppose for a moment that B is non-empty and recall that

$$t(B) = \sup\{t:\ t^{r+1} = q_0 t^r + q_1 t^{r-1} + \ldots + q_r \text{ for } (q_0,\ldots,q_r) \in B\}.$$

Definition 9.6

We shall say $p(\varphi)$ is the order of algorithm φ iff

$$(9.20)\quad p(\varphi) = \begin{cases} 0 & \text{if } B \text{ is empty,} \\ t(B) & \text{otherwise.} \end{cases}$$ ■

For $r = 0$, Definition 9.6 is equivalent to Definition 2.6 since $p(\varphi) = t(B) = \sup B$ for non-empty B. Note that in (9.19) we compare the value of $\varphi(x_0,\ldots,x_r,\mathfrak{N}(f,x_0,\ldots,x_r))$ to every solution element $S(\tilde{f}(x_0,\ldots,x_r))$ where $\tilde{f} \in V(f)$. The "classical" definition of the order is based on the comparison between $\varphi(x_0,\ldots,x_r,\mathfrak{N}(f,x_0,\ldots,x_r))$ and the solution $\alpha = S(f)$. We show in Appendix A that for all algorithms of practical interest the "classical" order is equal to $p(\varphi)$.

It is obvious that a result analogous to Theorem 9.3 can be stated for an algorithm φ. Namely if

$$(9.21)\quad \|\varphi(x_0,\ldots,x_r,\mathfrak{N}(f,x_0,\ldots,x_r))-S(\tilde{f}(x_0,\ldots,x_r))\| = O(\|x_0-\alpha\|^{p_0}\cdot\ldots\cdot\|x_r-\alpha\|^{p_r})$$

for $p_i \geq 0$ and $\sum_{i=0}^{r} p_i \geq 1$, then the order $p(\varphi) \geq p$ where p is the unique positive zero of $t^{r+1} - (p_0t^r + \ldots + p_r)$. Furthermore if (9.21) is sharp then $p(\varphi) = p$.

We now prove that the order of an algorithm can not exceed the order of information.

Theorem 9.4

For any algorithm φ, $\varphi \in \Phi(\mathfrak{N},S)$,

$$(9.22)\quad p(\varphi) \leq p(\mathfrak{N},S).$$ ■

<u>Proof</u>

Compare with the proof of Theorem 2.3. Without loss of generality we can assume $B \neq \emptyset$. Let $(q_0, q_1, \ldots, q_r)$ be any element of B. Let $f \in \mathfrak{J}_0$ and $\tilde{f} \in V(f)$. Then

$$\| S(\tilde{f}(x_0,\ldots,x_r)) - S(f) \| \le \| \varphi(x_0,\ldots,x_r,\mathfrak{N}(f,x_0,\ldots,x_r)) - S(\tilde{f}(x_0,\ldots,x_r)) \| +$$

$$\| \varphi(x_0,\ldots,x_r,\mathfrak{N}(f,x_0,\ldots,x_r)) - S(f) \| = o(\| x_0-\alpha\|^{q_0-\eta} \cdot \ldots \cdot \|x_r-\alpha\|^{q_r-\eta}), \ \forall \eta > 0.$$

This proves that $(q_0, q_1, \ldots, q_r) \in A$, see (9.10), and $t(B) \le t(A)$. This yields $p(\varphi) \le p(\mathfrak{N}, S)$. ■

The bound $p(\mathfrak{N},S)$ is achieved by the order of any interpolatory algorithm.

<u>Theorem 9.5</u>

For any interpolatory algorithm φ^I, $\varphi^I \in \Phi(\mathfrak{N},S)$,

(9.23) $p(\varphi^I) = p(\mathfrak{N},S)$. ■

<u>Proof</u>

The method of proof is similar to that used in Theorems 2.4 and 9.4. ■

Theorems 9.4 and 9.5 prove that any interpolatory algorithm achieves the maximal order $p(\mathfrak{N},S)$ in the class $\Phi(\mathfrak{N},S)$.

<u>Example 9.1</u> Nonlinear Equations

As in Example 6.3 define $S(f) = f^{-1}(0)$. Consider

(9.24) $\mathfrak{N}(f,x_0,\ldots,x_r) = [f(x_0),f'(x_0),\ldots,f^{(k)}(x_0),\ldots,f(x_r),f'(x_r),\ldots,f^{(k)}(x_r)]$

for a given $k \geq 0$. Let W be the class of analytic functions. Thus $\tilde{f} \in W$ means that $\tilde{f}(x,t)$ is analytic with respect to x and t.

Case 1, N = 1. Thus, f is a scalar function of a real or complex scalar variable. Then $\tilde{f} \in V(f)$ yields that $f(x,t) = f(t) + G(x,t) \prod_{j=0}^{r}(x-x_j)^{k+1}$ for an analytic function $G(x,t)$. It is easy to verify that the order of information $p(\mathfrak{N},S) = p(k,r)$ where $p(k,r)$ is the unique positive zero of the polynomial $t^{r+1}-(k+1)\sum_{j=0}^{r} t^j$ and $k+1 \leq p(k,r) < k+2$, $\lim_{r\to+\infty} p(k,r) = k+2$. It is well-known that the interpolatory algorithm $\varphi^I_{r,k}$ is now defined as follows:

(i) find an interpolatory polynomial ω of degree $\leq (k+1)(r+1)-1$ such that

$$\omega^{(i)}(x_j) = f^{(i)}(x_j), \quad i = 0,1,\ldots,k;\ j = 0,1,\ldots,r,$$

(ii) define $x_{r+1} = \varphi^I_{r,k}(x_0,\ldots,x_r,\mathfrak{N}(f,x_0,\ldots,x_r))$ as a zero of ω with a certain criterion of its choice (for instance the nearest zero to x_0).

For some values of k and r we get the known iterations. For example, $k = 0$ and $r = 1$ yields the secant iteration, $k = 1$ and $r = 0$ yields the Newton iteration. For a detailed discussion, see Traub [64].

Case 2, N ≥ 2. Thus, f is a multivariate or an abstract function. From Theorem 5 in Woźniakowski [74] follows

$$(9.25)\quad p(\mathfrak{N},S) = k+1, \quad \forall\, r.$$

This means that $p(\mathfrak{N},S)$ does not depend on r and the information contained in $f^{(i)}(x_j)$ for $i = 0,1,\ldots,k$ and $j = 1,2,\ldots,r$ does not help to increase the order.

However if a certain position of $x_1, x_2, \ldots, x_r$ is assumed than $p(\mathfrak{N}, S)$ can be larger than $k+1$ for $r \geq N+1$. Examples include multivariate secant iteration ($k = 0$, $r = N$), and the generalization of the interpolatory iteration $\varphi^I_{r,k}$. See for instance Brent [72], Jankowska [75], Ortega and Rheinboldt [70], Pleshakov [77] and Woźniakowski [74]. ∎

We briefly discuss the complexity of an algorithm φ which uses an information operator $\mathfrak{N}$ with memory. Suppose that for given initial approximations $x_0, x_1, \ldots, x_r$, the algorithm φ produces the sequence $\{x_i\}$ $x_{i+1} = \varphi(x_i, \ldots, x_{i-r}, \mathfrak{N}(f, x_i, \ldots, x_{i-r}))$ such that

$$(9.26) \quad e_i = G_i \, e_{i-1}^{p_0} \, e_{i-2}^{p_1} \cdot \ldots \cdot e_{i-r-1}^{p_r}, \quad e_i = \| x_i - \alpha \|, \quad i = r+1, \ldots$$

where $p_i \geq 0$ and $q = \sum_{i=0}^{r} p_i > 1$. Let p be the unique positive zero of $t^{r+1} - (p_0 t^r + \ldots + p_r)$. Assume that the constants $G_i = G_i(f)$ satisfy the relation

$$(9.27) \quad 0 < \underline{G} \leq G_i \leq \bar{G} < +\infty.$$

To simplify the complexity analysis we assume that there exist constants Γ, $\Gamma < 1$, C_1 and C_2 such that

$$(9.28) \quad C_1 \Gamma^{p^i} \leq e_i \leq C_2 \Gamma^{p^i} \quad \text{for } i = 1, 2, \ldots, r,$$

$$(9.29) \quad C_1^{1-q} < \underline{G}, \; C_2^{1-q} > \bar{G}.$$

Then, it is easy to verify that

$$(9.30) \quad C_1 \Gamma^{p^i} \leq e_i \leq C_2 \Gamma^{p^i}, \quad \forall \, i.$$

Recall that we want to find x_k such that k is the smallest integer for which $e_k \le \varepsilon' e_0$ for given $\varepsilon' \in (0,1)$. Let ε, $\varepsilon < \varepsilon'$, be such that $e_k = \varepsilon\, e_0$. From (9.30) we find

$$(9.31)\quad \frac{\log\left(\log\left(\frac{c_1}{\varepsilon e_0}\right)/\log\frac{1}{\Gamma}\right)}{\log p} \le k \le \frac{\log\left(\log\left(\frac{c_2}{\varepsilon e_0}\right)/\log\frac{1}{\Gamma}\right)}{\log p}$$

For small ε, $k \cong (\log\log 1/\varepsilon)/\log p$. Let $\mathrm{comp}(\varphi,f)$ be the complexity of computing x_k from $x_0, x_1, \ldots, x_r$. Assume that the cost of every iterative step is equal to $c(\varphi,f)$. Then

$$(9.32)\quad \mathrm{comp}(\varphi,f) = k\, c(\varphi,f) = z(\varphi,f)\, g(k,\varepsilon)$$

where $g(k,\varepsilon) = k \log p = (\log\log 1/\varepsilon)(1 + O(\varepsilon))$ and

$$(9.33)\quad z(\varphi,f) = \frac{c(\varphi,f)}{\log p}$$

is the <u>complexity index</u> of φ for f.

We discuss the cost of one iterative step. To perform the ith step we have to compute $\mathfrak{N}(f,x_{i-1},x_{i-2},\ldots,x_{i-r-1})$ and next $\varphi(x_{i-1},\ldots,x_{i-r-1},\mathfrak{N}(f,x_{i-1},\ldots,x_{i-}$
Assume that

$$(9.34)\quad \mathfrak{N}(f,x_{i-1},\ldots,x_{i-r-1}) = [\bar{\mathfrak{N}}(f,x_{i-1}),\bar{\mathfrak{N}}(f,x_{i-2}),\ldots,\bar{\mathfrak{N}}(f,x_{i-r-1})]$$

for a certain permissible information operator $\bar{\mathfrak{N}}$ without memory. Thus the information complexity $\mathrm{comp}(\mathfrak{N}(f,x_{i-1},\ldots,x_{i-r-1})) = \mathrm{comp}(\bar{\mathfrak{N}}(f,x_{i-1}))$ since we reuse the previously computed information $\bar{\mathfrak{N}}(f,x_{i-2}),\ldots,\bar{\mathfrak{N}}(f,x_{i-r-1})$. This

means that the information complexity does not depend on r, the size of memory. Let $\mathrm{comp}(\mathfrak{N}) \equiv \mathrm{comp}(\mathfrak{N}(f,x_i))$.

Let $\mathrm{comb}(\varphi,r)$ denote the combinatory complexity of an algorithm φ. Then

$$(9.35)\quad z(\varphi,f) = \frac{\mathrm{comp}(\mathfrak{N}) + \mathrm{comb}(\varphi,r)}{\log p}\,.$$

We seek an algorithm with the minimal complexity index. Since the order of an algorithm φ is no larger than the order of information $p(\mathfrak{N},S)$ and there exist algorithms φ with $p(\varphi) = p(\mathfrak{N},S)$ we get

$$(9.36)\quad \frac{\mathrm{comp}(\mathfrak{N}) + \underline{\mathrm{comb}}(\mathfrak{N},r)}{\log p(\mathfrak{N},S)} \le \inf_{\varphi} z(\varphi,f) \le \frac{\mathrm{comp}(\mathfrak{N}) + \overline{\mathrm{comb}}(\mathfrak{N},r)}{\log p(\mathfrak{N},S)}$$

where $\underline{\mathrm{comb}}(\mathfrak{N},r)$ is the minimal combinatory complexity of algorithms which use information $\mathfrak{N}$ and have order greater than unity, $\overline{\mathrm{comb}}(\mathfrak{N},r)$ is the minimal combinatory complexity of algorithms with the maximal order $p(\mathfrak{N},S)$. If $\mathrm{comp}(\mathfrak{N}) \gg \overline{\mathrm{comb}}(\mathfrak{N},r)$ then

$$(9.37)\quad \inf_{\varphi} z(\varphi,f) \cong \frac{\mathrm{comp}(\mathfrak{N})}{\log p(\mathfrak{N},S)}\,.$$

Note, however, that the inequality $\mathrm{comp}(\mathfrak{N}) \gg \overline{\mathrm{comb}}(\mathfrak{N},r)$ can usually hold only for small r or for "sufficiently hard-to-compute" f.

10. EXTENSIONS AND OPEN PROBLEMS

We conclude this paper by a partial list of extensions and open problems which will be studied in the future.

1. We show in this paper that if the index of a problem is infinite the problem cannot be solved iteratively by a one-point linear information operator with finite cardinality. What is the characterization of all problems with finite index? Furthermore, what is the characterization of all problems with index(S) = s for a given integer s? We are also interested in the same question for the mth index of a linear information operator with or without memory.

We proved that finite dimensional nonlinear equations $S(f) = f^{-1}(0)$ can be solved by *iterative algorithms of finite order* using linear information operators with finite cardinality. Is that the most general form of S which can be solved iteratively? The following example shows this not to be the case.

<u>Example 10.1</u>

Let $f : [0,1] \to R$ be a smooth scalar function. Define

$$[F(f)](t) = [Af](t) + d(t,f(t)), \qquad t \in [0,1],$$

where A is a linear operator and d is a smooth function of two variables. Let $\mathfrak{F}_0$ be a class of functions f for which the equation $[F(f)](t) = 0$ has a unique simple zero in [0,1]. Define the problem

$$(10.1)\quad S(f) = g([F(f)]^{-1}(0))$$

where g is a smooth one-to-one function and its inverse g^{-1} is a Lipschitz function. Thus, (10.1) means find the solution β of the equation $[F(f)](t) = 0$ and then compute $\alpha = g(\beta)$. Consider the following one-point linear information operator

$$\mathfrak{N}(f,x) = [[Af](y),[Af]'(y),f(y),f'(y)] \text{ where } y = g^{-1}(x).$$

Note that $\text{card}(\mathfrak{N}) \le 4$. Knowing $\mathfrak{N}(f,x)$ we can compute $[F(f)](y) = [Af](y) + d(y,f(y))$ and $[F(f)]'(y) = [Af]'(y) + \partial_1 d(y,f(y)) + \partial_2 d(y,f(y))\, f'(y)$ and then apply Newton iteration to approximate $\beta = g^{-1}(\alpha)$. (∂_i denotes the partial derivative with respect to the ith argument.) Consider the algorithm

$$\varphi(x,\mathfrak{N}(f,x)) = g(y-[F(f)](y)/[F(f)]'(y)).$$

Then

$$\begin{aligned}\varphi(x,\mathfrak{N}(f,x))-\alpha &= g(\beta+O((y-\beta)^2)) - g(\beta) = O((g^{-1}(x)-g^{-1}(\alpha))^2)\\ &= O((x-\alpha)^2).\end{aligned}$$

This proves that the order of φ and the order of information are at least equal to 2. Thus, problem (10.1) can be solved by iteration. As a particular example set $Af = f^{(j)}$, $d \equiv 0$ and $g(x) = x$. Then $\alpha = S(f)$ is the unique simple solution of the equation $f^{(j)}(t) = 0$. ∎

We wish to find the most general form of S which can be solved iteratively and propose the following conjecture.

Let D_0 and D_1 be open subsets of $\mathbb{C}^N$, $1 \le N < +\infty$. Let $\mathfrak{J}_1$ denote the class of functions $f : D_0 \to D_1$.

<u>Conjecture 10.1</u>

If the problem S can be solved by iterative algorithms using one-point linear information operators with finite cardinality then there exists an

an operator $F : D_F \subset \mathfrak{I}_1 \to \mathfrak{I}_1$ and a function $g : D_g \subset \mathbb{C}^N \to \mathfrak{I}_2$ such that

$$(10.2)\quad S(f) = g([F(f)]^{-1}(0)), \qquad \forall f \in \mathfrak{I}_0. \qquad \blacksquare$$

Conjecture 10.1 states that essentially only nonlinear equations can be solved by iteration. Indeed, (10.2) means that $\alpha = S(f)$ is the transformed value of β, $\alpha = g(\beta)$, where β is a solution of the transformed nonlinear equation $[F(f)](t) = 0$. Note that in Conjecture 10.1 we do not specify properties of F and g. We merely assume their existence. Example 10.1 provides an example of F and g such that the problem $g([F(f)]^{-1}(0))$ can be solved iteratively. It would be interesting to find the most general form of F and g which permits a problem to be iteratively solved.

In Part A we showed many apparently diverse problems could be handled within the same general framework. However if Conjecture 10.1 is true then essentially only problems that we already knew could be solved by iteration are included within the iterative information model of this paper.

2. We discuss the classification of more general linear information operators than those considered in this paper. These operators are also of practical and theoretical interest. We define an iterative linear information operator as follows. Let

$$(10.3)\quad \mathfrak{N}(f,x_0,\ldots,x_r) =$$
$$[L_1(f,\xi_1(x_0)),\ldots,L_n(f,\xi_n(x_0)),\ldots,L_1(f,\xi_1(x_r)),\ldots,L_n(f,\xi_n(x_r))]^t$$

where

$$\xi_1(x) = x$$
$$\xi_{j+1}(x) = \xi_{j+1}(x,L_1(f,\xi_1(x)),\ldots,L_j(f,\xi_j(x))), \quad j = 1,2,\ldots,n-1$$

and $L_j(f,x)$ is a linear functional with respect to f. Thus $L_j(f,\xi_j(x))$

depends on the previously computed information. The parameter r measures the size of "memory". For $r = 0$, $\mathfrak{N}$ is an <u>iterative linear information operator without memory</u>. For $r \geq 1$, $\mathfrak{N}$ reuses the previously computed information at $x_1,\ldots,x_r$ and is an <u>iterative linear information operator with memory</u>. For $\xi_j(x) \equiv x$, $j = 1,2,\ldots,n$ we get a <u>one-point iterative linear information operator</u> which for $r = 0$ was considered in Sections 4 to 8. If there exists j such that $\xi_j(x) \not\equiv x$, $\mathfrak{N}$ is a <u>multipoint iterative linear information operator</u> since the information $L_1(f,\xi_1(x)),\ldots,L_n(f,\xi_n(x))$ is computed at least at two different points. Examples of multipoint iterations for nonlinear equations may be found in Brent [76], Kacewicz [75], Kung and Traub [74], Meersman [76a, 76b] and Woźniakowski [76].

This classification is schematized in Figure 2.

	$r = 0$	$r \geq 1$
$\forall j,\ \xi_j(x) \equiv x$	one-point without memory	one-point with memory
$\exists j,\ \xi_j(x) \not\equiv x$	multipoint without memory	multipoint with memory

Figure 2

<u>Remark 10.1</u>

For nonlinear iterative information operators there is no difference between one-point and multipoint operators. For the nonlinear case we can distinguish iterative information operators without memory $\mathfrak{N} = \mathfrak{N}(f,x_0)$ which are considered in Section 2 and iterative information operators with memory which are considered in Section 9. ■

It would be of interest to generalize the results of Sections 4 to 8 to multipoint iterative linear information operators with or without memory. In particular, we are interested in the questions raised in extension 1 to multipoint iterations or iterations with memory. We would also like to extend the concept and properties of basic linear information operator (see Sections 6 and 7). That is, given a number m, find a linear information operator $\mathfrak{N}^*$ with minimal cardinality of order at least m. Does the conclusion of Theorem 7.3 continue to hold?

3. What is the minimal number of linear functionals in (10.3) to iteratively solve the system of nonlinear equations $f(x) = 0$, $f: D \subset \mathbb{C}^N \to \mathbb{C}^N$? Newton iteration shows that $O(N^2)$ linear functionals are sufficient. It follows from Lemma 4.3 and Theorem 4.2 in Traub and Woźniakowski [76b] that any iteration based on a one-point linear information operator requires at least the evaluation of f and f'. This holds even if $N = \infty$. This result is related to informational requirements of convergent price mechanisms in mathematical economics. See Saari and Simon [76].

Kacewicz [77] conjectures that $N + c N^2$ linear functionals are needed (without the restriction to one-point iteration), where c is a positive constant. Kacewicz has obtained partial results on the minimal number of linear functionals for $N = 1$ and 2.

4. In order to derive lower bounds on complexity we require upper bounds on the order of information for fixed information. In particular, let f be a scalar nonlinear function with a simple zero and let $S = f^{-1}(0)$. Kung and Traub [74] show there exists a multipoint linear information operator using the linear functionals

(10.4) $L_j(f,x) \equiv f^{(k_j)}(x), \ k_j \geq 0, \ \ j = 1,2,\ldots,n,$

such that

(10.5) $p(\mathfrak{N},S) = 2^{n-1}$.

They conjecture that

(10.6) $p(\mathfrak{N},S) \leq 2^{n-1}$

for all linear multipoint information operators which use the functionals of the form (10.4). This conjecture was established for n = 1,2 by Kung and Traub [76b], for n = 3 by Meersman [76a, 76b] and for "Hermite" information with arbitrary n by Woźniakowski [76]. Wasilkowski [77] proves this conjecture holds whenever the information operator is well-poised in the sense of Birkhoff complex interpolation.

We generalize the Kung-Traub conjecture.

Conjecture 10.2

Let f be any non-linear problem with a simple zero and let $S = f^{-1}(0)$. Let $L_1,\ldots,L_n$ be arbitrary linear functionals and let $\xi_1,\ldots,\xi_n$ be arbitrary functions. Then

$$p(\mathfrak{N},S) \leq 2^{n-1} \quad \text{for } r = 0,$$
$$p(\mathfrak{N},S) < 2^{n} \quad \text{for } 0 < r < \infty .$$

■

See also Kacewicz and Woźniakowski [77] where the maximal order of information for multipoint iterations is discussed.

11. COMPARISON OF RESULTS FROM GENERAL AND ITERATIVE INFORMATION MODELS

For the reader's convenience we compare some of the results from the general information model of Part A and the iterative information model of Part B. See Figures 3 and 4 for a summary.

Part B Iterative Information Model

	$p(\mathfrak{N},S) = 0$	$p(\mathfrak{N},S) = 1$	$1 < p(\mathfrak{N},S) < \infty$	$p(\mathfrak{N},S) = \infty$
$d(\mathfrak{N},S)$	> 0	0	0	0
$\text{comp}(\mathfrak{N},S,\epsilon)$	undefined	$\simeq \lg(1/\epsilon)$	$\simeq \lg\lg(1/\epsilon)$	constant

Figure 3

Part A General Information Model

	$r(\mathfrak{N},S) = 0$	$0 < r(\mathfrak{N},S) \le \epsilon$	$r(\mathfrak{N},S) > \epsilon$
$p(\mathfrak{N},S)$	∞	0	0
$\text{comp}(\mathfrak{N},S,\epsilon)$	constant	"any" monotonically increasing function	undefined

Figure 4

We comment on these Figures. In both Figures we give the asymptotic dependence of $\text{comp}(\mathfrak{N},S,\epsilon)$ as a function of ϵ. If the problem cannot be solved to within ϵ, we say the complexity is undefined. In the iterative information model, the order of information, $p(\mathfrak{N},S)$, is basic. Thus if $p(\mathfrak{N},S) = 0$, the diameter of information is positive and the complexity is

undefined. In the general information model, the radius of information, $r(\mathfrak{N},S)$, is basic. Thus if $r(\mathfrak{N},S) = 0$, then the order is infinite and the complexity is independent of ϵ.

Order of information was not defined in Part A. However the definition of Part B can be used if we recognize that in Part A, $\mathfrak{N}(f,x)$ is independent of x. Then the order of information must be either zero or infinite.

APPENDIX A

We discuss the relationship between the order $p(\varphi)$ of an algorithm φ defined by (2.21) or (9.20) and the "classical" definition of order which can be stated as follows. Let $p_0,\dots,p_r$ be real numbers such that $p_i = 0$ or $p_i \geq 1$ and $q = \sum_{i=0}^{r} p_i > 1$. Suppose that φ uses the information operator $\mathfrak{N}(f,x_0,\dots,x_r)$ for $f \in \mathfrak{J}_0$. Assume that for every $f \in \mathfrak{J}_0$, $\alpha = S(f)$, there exist $\Gamma = \Gamma(f) > 0$ and $c = c(f) < +\infty$ such that

$$(A.1)\quad \|\varphi(x_0,\dots,x_r;\mathfrak{N}(f,x_0,\dots,x_r))-\alpha\| \leq c(f)\,\|x_0-\alpha\|^{p_0}\|x_1-\alpha\|^{p_1}\cdot\ldots\cdot\|x_r-\alpha\|^{p_r}$$

for all $\|x_0-\alpha\| \leq \|x_1-\alpha\| \leq \dots \leq \|x_r-\alpha\| \leq \Gamma(f)$.

Assume that (A.1) is sharp, i.e., there exists $f_0 \in \mathfrak{J}_0$, $\alpha_0 = S(f_0)$, such that

$$(A.2)\quad \|\varphi(x_0,\dots,x_r,\mathfrak{N}(x_0,\dots,x_r,f_0))-\alpha_0\| \geq c_0(f_0)\,\|x_0-\alpha_0\|^{p_0}\|x_1-\alpha_0\|^{p_1}\cdot\ldots\cdot\|x_r-\alpha_0\|^{p_r}$$

for all $\|x_0-\alpha_0\| \leq \dots \leq \|x_r-\alpha_0\| \leq \Gamma(f_0)$ and $c_0(f_0) > 0$. Then we shall call the unique positive zero p of the polynomial $t^{r+1} - (p_0t^r + \dots + p_r)$ the <u>"classical" order of φ</u>. (See among others Traub [64].) Note that for $r = 0$, i.e , $\mathfrak{N}$ is an information operator without memory, $p = p_0$.

It is easy to verify that if

$$(A.3)\quad c(f)\,\Gamma(f)^{q-1} < 1$$

and all initial approximations $x_0, x_1,\dots,x_r$ satisfy

$$(A.4)\quad e_i = \|x_i-\alpha\| \leq \Gamma(f) \quad \text{for } i = 0,1,\dots,r$$

then φ generates the sequence $\{x_i\}$, $x_i = \varphi(x_i,\ldots,x_{i-r},\mathfrak{N}(f,x_i,\ldots,x_{i-r}))$ for $i \geq r+1$ which has the property

$$\text{(A.5)} \quad \lim_i x_i = \alpha,$$

$$\text{(A.6)} \quad e_i = O(\zeta^{p^i})$$

for a certain $\zeta < 1$.

The number $\Gamma(f)$ can be interpreted as the radius of a ball of convergence since φ convergences for all initial approximations from $J = \{x: \|x-\alpha\| \leq \Gamma(f)\}$. The constant $c(f)$ can be interpreted as the "asymptotic constant" which satisfies

$$\text{(A.7)} \quad \overline{\lim_i} \frac{e_{i+1}}{e_i^{p_0} \cdot \ldots \cdot e_{i-r}^{p_r}} \leq c(f).$$

We are ready to prove

<u>Lemma A.1</u>

Let $\mathfrak{N}$ be convergent information. Suppose that (A.1) and (A.2) hold and additionally

$$\text{(A.8)} \quad \Gamma_0 = \underline{\lim}_{x_0,\ldots,x_r \to \alpha} \Gamma(\tilde{f}(x_0,x_1,\ldots,x_r)) > 0,$$

$$\text{(A.9)} \quad c_0 = \overline{\lim}_{x_0,\ldots,x_r \to \alpha} c(\tilde{f}(x_0,\ldots,x_r)) < +\infty$$

for all $f \in \mathfrak{F}_0$ and all $\tilde{f} \in V(f)$. Then $p(\varphi) = p$. ∎

Proof

Since $\mathfrak{N}$ is convergent, $\tilde{\alpha} = S(\tilde{f}(x_0,x_1,\ldots,x_r))$ tends to $\alpha = S(f)$ as $x_0,x_1,\ldots,x_r$ approach α for every $\tilde{f} \in V(f)$. From (A.8) we get

$$\| x_i-\tilde{\alpha}\| \le \Gamma(\tilde{f}(x_0,\ldots,x_r)), \quad i = 0,1,\ldots,r,$$

for sufficiently small $\max_{0\le i\le r} \| x_i-\alpha\|$. This means that $x_0,\ldots,x_r$ can be treated as approximations to α and $\tilde{\alpha}$. Let $x = \varphi(x_0,\ldots,x_r,\mathfrak{N}(f,x_0,\ldots,x_r))$. From (A.1) we have

$$\text{(A.10)} \quad \|\alpha-\tilde{\alpha}\| \le \|x-\alpha\| + \|x-\tilde{\alpha}\| \le c(f) \prod_{i=0}^{r} \|x_i-\alpha\|^{p_i} + c(\tilde{f}(x_0,\ldots,x_r)) \prod_{i=0}^{r} \|x_i-\tilde{\alpha}\|^{p_i}.$$

Choose j such that $p_j \ge 1$. Since $\sum_{i=0}^{r} p_i > 1$ and $c(\tilde{f}(x_0,\ldots,x_r))$ is bounded due to (A.9) then (A.10) can be rewritten

$$\| \alpha-\tilde{\alpha}\| = o(\| x_j-\alpha\|) + o(\| x_j-\alpha\| + \| \alpha-\tilde{\alpha}\|).$$

This yields $\| \alpha-\tilde{\alpha}\| = o(\| x_j-\alpha\|)$. From this and (A.10) we get

$$\| \alpha-\tilde{\alpha}\| = O(\prod_{i=0}^{r} \| x_i-\alpha\|^{p_i}).$$

This proves that $(p_0,p_1,\ldots,p_r) \in B$, see (9.19), and consequently $p \le p(\varphi)$.

Set $\tilde{f}(x_0,\ldots,x_r) = f_0$ where f_0 satisfies (A.2). Let $(q_0,q_1,\ldots,q_r) \in B$. Then

$$\lim_{\| x_0-\alpha\|\le\ldots\le\| x_r-\alpha\|\to 0} \frac{\| \varphi(x_0,\ldots,x_r,\mathfrak{N}(f_0,x_0,\ldots,x_r))-\alpha\|}{\| x_0-\alpha\|^{q_0-\eta} \cdot\ldots\cdot \| x_r-\alpha\|^{q_r-\eta}} = 0, \quad \forall\, \eta > 0.$$

From (A.2) it easily follows that

$$q_0 + q_1 + \ldots + q_i \le p_0 + p_1 + \ldots + p_i \quad \text{for } i = 0,1,\ldots,r.$$

From Lemma 9.1 we get $t \in (0,p]$ where $t^{r+1} = q_0 t^r + \ldots + q_r$. Since $(q_0,\ldots,q_r)$ is any element of B, $p(\varphi) = t(B) \le p$. Thus $p(\varphi) = p$ which completes the proof. ∎

Equations (A.8) and (A.9) state that the radius $\Gamma(\tilde{f}(x_0,\ldots,x_r))$ of the problem element $\tilde{f}(x_0,\ldots,x_r)$ is bounded from below roughly by $\Gamma_0 > 0$ and the asymptotic constant $c(\tilde{f}(x_0,\ldots,x_r))$ is bounded from above roughly by $c_0 < +\infty$. The assumptions (A.8) and (A.9) hold for all algorithms of practical interest since $\Gamma(\tilde{f}(x_0,\ldots,x_r))$ and $c(\tilde{f}(x_0,\ldots,x_r))$ are continuous with respect to $x_0,\ldots,x_r$ and $\Gamma_0 = \Gamma(\tilde{f}(\alpha,\ldots,\alpha)) > 0$, $c_0 = c(\tilde{f}(\alpha,\ldots,\alpha)) < +\infty$. Therefore for all practical cases, the "new" definition of order coincides with the "classical" one, $p(\varphi) = p$.

ACKNOWLEDGMENTS

We thank G. Wasilkowski and A. Werschulz for their comments on the manuscript.

GLOSSARY

We list basic concepts used throughout the paper. We mention a symbol, its meaning and section reference where the symbol appears for the first time.

Symbol	Meaning	Section Reference
S	the solution operator, sometimes called the problem, $S: \mathfrak{I}_0 \to \mathfrak{I}_2$, $\mathfrak{I}_0 \subset \mathfrak{I}_1$	2, (2.1)
$\mathfrak{I}_0$	the domain of S	2
$\mathfrak{I}_1$	linear space, $\mathfrak{I}_0 \subset \mathfrak{I}_1$	2
$\mathfrak{I}_2$	the range of S	2
α	the solution element, $\alpha = S(f)$	2
f	the problem element, $f \in \mathfrak{I}_0$	2
ϵ'	error parameter	2
x_0	a given initial approximation	2
$y = y(f)$	an ϵ'-approximation, $\\|y(f)-\alpha\\| \le \epsilon' \\|x_0-\alpha\\|$	2
$\mathfrak{N}$	the iterative information operator, $\mathfrak{N}: D_{\mathfrak{N}} \to \mathfrak{I}_3$	2
$\mathfrak{I}_3$	the range of $\mathfrak{N}$	2
$\tilde{f}$	a function, $\tilde{f}: D_{\tilde{f}} \to \mathfrak{I}_0$, $f \in W$ and $\mathfrak{N}(\tilde{f}(x),x) = \mathfrak{N}(f,\tilde{x})$	2, def. 2.1
W	the regularity space	2, def. 2.1
$V(f)$	the set of functions $\tilde{f}$	2, def. 2.1
$d(\mathfrak{N},S)$	the limiting diameter of information $\mathfrak{N}$ for the problem S	2, def. 2.2
φ	an algorithm, $\varphi: D_\varphi \subset \mathfrak{I}_2 \times \mathfrak{N}(D_{\mathfrak{N}}) \to \mathfrak{I}_2$	2, (2.11)
$\Phi(\mathfrak{N},S)$	the class of all algorithms using the information $\mathfrak{N}$ for the problem S	2
$e(\varphi)$	the limiting error of algorithm φ	2, (2.13)

Symbol	Meaning	Section Reference
φ^I	an interpolatory algorithm	2, (2.15)
$p(\mathfrak{N},S)$	the order of information $\mathfrak{N}$ for the problem S	2, (2.18)
$p(\varphi)$	the order of algorithm φ	2, (2.21)
P	the set of primitives	3
$comp(\mathfrak{N}(f,x))$	the information complexity for computing $\mathfrak{N}(f,x)$ where $\mathfrak{N}$ is a permissible information operator	3
$comp(\varphi(x,\mathfrak{N}(f,x)))$	the combinatory complexity for computing $\varphi(x,\mathfrak{N}(f,x))$ where φ is a permissible algorithm	3
$z(\varphi,f)$	the complexity index of φ for f	3, (3.10)
$z(\varphi)$	the complexity index of algorithm φ	3, (3.11)
$z(\mathfrak{N},S)$	the complexity index of the information $\mathfrak{N}$ for problem S	3, (3.15)
$\Phi_{perm}(\mathfrak{N},S)$	the class of all permissible algorithms	3
φ^{mc}	a minimal complexity index algorithm	3, (3.16)
$comp(\mathfrak{N})$	the information complexity	3, (3.17)
$\mathfrak{N}^o$	an optimal information operator in the class Ψ	3, (3.22)
$\mathfrak{N}_1 \subset \mathfrak{N}_2$	$\ker \mathfrak{N}_2(\cdot,x) \subset \ker \mathfrak{N}_1(\cdot,x), \ \forall x$	4, def. 4.1
$\mathfrak{N}_1 \asymp \mathfrak{N}_2$	$\ker \mathfrak{N}_1(\cdot,x) = \ker \mathfrak{N}_1(\cdot,x), \ \forall x$	4, def. 4.1
$A^\perp$	algebraic complement of A	4, def. 4.4
codim A	codimension of A	4, def. 4.5
$card(\mathfrak{N})$	the cardinality of the information $\mathfrak{N}$	4, def. 4.2
Lip(k)	the class of k times differentiable functions whose kth derivatives satisfy a Lipschitz condition	4, (4.15)
$\mathfrak{N} \in Lip(k)$	$\mathfrak{N}$ belongs to the class Lip(k)	4, def. 4.3
$IT(\mathfrak{N},S)$	the class of all iterative algorithms using the information $\mathfrak{N}$ for the problem S	5, def. 5.1
index(S)	the index of the problem S	6, def. 6.1
$\mathfrak{N}^*$	a basic linear information operator	6, (6.7)
$S \in Lip(0)$	the solution operator S belongs to the class Lip(0)	6, def. 6.2

Symbol	Meaning	Section Reference
index(S,m)	the mth index of the problem S	7, def. 7.1
$\mathfrak{N}_m^*$	an mth basic linear information operator	7, (7.10)
$S \in \mathrm{Lip}(k(m))$	the solution operator S belongs to the class Lip(k(m))	7, def. 7.2
p(n,S)	the nth maximal order for the problem S	8, def. 8.1
$\mathfrak{N}^{mo}$	an nth maximal order information for the problem S	8, def. 8.1
z(n,S)	the nth minimal complexity index for the problem S	8, def. 8.2
$n^*(S)$	the optimal cardinality number for the problem S	8, def. 8.2
$\mathfrak{N}^{oi}$	an optimal information operator for the problem S	8, def. 8.2

BIBLIOGRAPHY

Brent [72] — Brent, R.P , "The Computational Complexity of Iterative Methods for Systems of Nonlinear Equations," Proc. Complexity Symposium, edited by Miller and Thatcher, Plenum Press, New York 1972, 61-71.

Brent [76] — Brent, R.P., "A Class of Optimal-Order Zero-Finding Methods Using Derivative Evaluations," in Analytic Computational Complexity, edited by J.F Traub, Academic Press, 1976, 59-75.

Brent, Winograd and Wolfe [73] — Brent, R.P., Winograd, S. and Wolfe, P., "Optimal Iterative Processes for Root-finding," Num. Math. 20, 1973, 327-341.

Jankowska [75] — Jankowska, J., "Multivariate Secant Method," Ph.D. Thesis, University of Warsaw, to appear in SIAM J. Num. Anal.

Kacewicz [75] — Kacewicz, B., "Integrals with a Kernel in the Solution of Nonlinear Equations in N Dimensions," Dept. of Computer Science Report, Carnegie-Mellon University, 1975.

Kacewicz [76a] — Kacewicz, B., "An Integral-Interpolation Iterative Method for the Solution of Scalar Equations," Numer. Math. 26, 1976, 355-365.

Kacewicz [76b] — Kacewicz, B., "The Use of Integrals in the Solution of Nonlinear Equations in N Dimensions," in Analytic Computational Complexity, Academic Press, 1976, 127-141.

Kacewicz [77] — Kacewicz, B., private communication.

Kacewicz and Woźniakowski [77] — Kacewicz, B. and Woźniakowski, H., "A Survey of Recent Problems and Results in Analytic Computational Complexity," Proc. of Mathematical Foundations of Computer Science 77, Springer Verlag, 53.

Knuth [76] — Knuth, D. E., "Big Omicron and Big Omega and Big Theta," SIGACT News, April 1976, 18-24.

Kung [76] — Kung, H. T., "The Complexity of Obtaining Starting Points for Solving Operator Equations by Newton's Method," in Analytic Computational Complexity, 1976, 35-57.

Kung and Traub [74] — Kung, H. T. and Traub, J. F., "Optimal Order of One-Point and Multipoint Iterations," J.ACM 21, 1974, 643-651.

Kung and Traub [76a] — Kung, H. T. and Traub, J. T., "All Algebraic Functions Can be Computed Fast," Dept. of Computer Science Report. Carnegie-Mellon University, 1976, J.ACM 25, 1978, 245-260.

Kung and Traub [76b] — Kung, H. T. and Traub, J. T., "Optimal Order and Efficiency for Iterations with Two Evaluations," SIAM J. Numer. Anal. 13, 1976, 84-99.

Meersman [76a] — Meersman, R., "Optimal Use of Information in Certain Iterative Processes," in Analytic Computational Complexity, Academic Press, 1976, 127-141.

Meersman [76b] — Meersman, R., "On Maximal Order of Families of Iterations for Nonlinear Equations," Doctoral Thesis, Vrije Universiteit Brussels, Brussels, 1976.

Ortega and Rheinboldt [70] — Ortega, J. M. and Rheinboldt, W. C., Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

Pleshakov [77] — Pleshakov, G. N., "On Efficiency of the Multidimensional Interpolation Iterations," in Russian, USSR Computational Math. and Math. Physics 17, 1977, 1153-1160.

Saari and Simon [76] — Saari, D. G. and Simon, C. P., "Effective Price Mechanisms," Dept. of Mathematics Report, University of Michigan, 1976.

Traub [61] — Traub, J. F., "On Functional Iteration and Calculation of Roots," Proc. 16th National ACM Conference, Los Angeles, CA, 1961.

Traub [64] — Traub, J. F., Iterative Methods for Solution of Equations, Prentice-Hall, Englewood Cliffs, NJ, 1964.

Traub and Woźniakowski [76a] — Traub, J. F. and Woźniakowski, H., "Strict Lower and Upper Bounds on Iterative Computational Complexity," in Analytic Computational Complexity, Academic Press, 1976, 15-34.

Traub and Woźniakowski [76b] — Traub, J. F. and Woźniakowski, H., "Optimal Linear Information for the Solution of Non-Linear Equations," in Algorithms and Complexity: New Directions and Recent Results, edited by J. F. Traub, Academic Press, 103-119.

Traub and Woźniakowski [76c] Traub, J. F. and Woźniakowski, H., "Optimal Radius of Convergence of Interpolatory Iterations for Operator Equations," Dept. of Computer Science Report, Carnegie-Mellon University, 1976. To appear in Aequationes Mathematica

Traub and Woźniakowski [77a] Traub, J. F. and Woźniakowski, H., "Convergence and Complexity of Newton Iteration for Operator Equations," Dept. of Computer Science Report, Carnegie-Mellon University, 1977. To appear in J.ACM.

Traub and Woźniakowski [77b] Traub, J. F. and Woźniakowski, H., "Convergence and Complexity of Interpolatory-Newton Iteration in a Banach Space," Dept. of Computer Science Report, Carnegie-Mellon University, 1977.

Traub and Woźniakowski [77c] Traub, J. F. and Woźniakowski, H., "General Theory of Optimal Error Algorithms and Analytic Complexity, Part A: General Information Model," Dept. of Computer Science Report, Carnegie-Mellon University, 1977.

Wasilkowski [77] Wasilkowski, G., "N-Evaluation Conjecture for Multipoint Iterations for the Solution of Scalar Nonlinear Equations," Master's Thesis, University of Warsaw, 1977.

Werschulz [77a] Werschulz, A. G., "Maximal Order and Order of Information for Numerical Quadrature," Mathematics Research Report 77-2, University of Maryland Baltimore County, 1977.

Werschulz [77b] Werschulz, A. G., "Optimal Order for Approximation of Derivatives," Mathematics Research Report 77-8, University of Maryland Baltimore County, 1977.

Woźniakowski [72] Woźniakowski, H., "On Nonlinear Iterative Processes in Numerical Methods," Doctoral Thesis, University of Warsaw, 1972 (in Polish).

Woźniakowski [74] Woźniakowski, H., "Maximal Stationary Iterative Methods for the Solution of Operator Equations," SIAM J. Numer. Anal., 11, 1974, 934-949.

Woźniakowski [75] | Woźniakowski, H., "Generalized Information and Maximal Order of Information for Operator Equations," SIAM J. Num. Anal. 12, 1975, 121-135.

Woźniakowski [76] | Woźniakowski, H., "Maximal Order of Multipoint Iterations Using n Evaluations," in Analytic Computational Complexity, Academic Press, 1976, 75-107.

AD-A063 757 CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER --ETC F/G 12/1
GENERAL THEORY OF OPTIMAL ERROR ALGORITHMS AND ANALYTIC COMPLEX--ETC(U)
NOV 78 J F TRAUB, H WOZNIAKOWSKI N00014-76-C-0370
UNCLASSIFIED CMU-CS-78-149 NL

2 OF 2
AD
A063757

END
DATE
FILMED
3-79
DDC

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CMU-CS-78-149	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) PART B. ITERATIVE INFO MODEL GENERAL THEORY OF OPTIMAL ERROR ALGORITHMS AND ANALYTIC COMPLEXITY		5. TYPE OF REPORT & PERIOD COVERED Interim 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) J.F. Traub and H. Wozniakowski		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0370
9. PERFORMING ORGANIZATION NAME AND ADDRESS Carnegie-Mellon University Computer Science Dept Pittsburgh, PA 15213		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		12. REPORT DATE November 1978 13. NUMBER OF PAGES 97
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Same as above		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601 |

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)